Taylor & Francis
Taylor & Francis Group

MITOGENOME ANNOUNCEMENT

🔓 OPEN ACCESS | Check for updates

The complete chloroplast genome of Korean endemic species, *Cirsium rhinoceros* (H.Lév. & vaniot) Nakai (Asteraceae)

Sang June Nam[a,b], Jung-Min Kim[c], Yongsung Kim[d,e] 🆔, Ja-Jung Ku[f], Su-Young Jung[g], You-Mi Lee[g], Seong-Ki Kim[h], Hong Xi[d,e], Jin-Young Song[a] and Jongsun Park[d,e] 🆔

[a]Research & Business Development Institute, Agricultural Corporation, Jeju Chunji, Republic of Korea; [b]Department of Agricultural Biotechnology, College of Agriculture and Life Sciences, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul; [c]Gyeongsangnam-do Geumwonsan Mountain Forest Resource Management Office, Gyeongsangnam-do, Republic of Korea; [d]InfoBoss Co., Ltd., Seoul, Republic of Korea; [e]InfoBoss Research Center, Seoul, Republic of Korea; [f]Forest Bioinformation Division, National Institute of Forest Science, Suwon, Republic of Korea; [g]Korea National Arboretum, Pocheon-si, Republic of Korea; [h]Department of Life science, Chung-Ang University, Seoul, Republic of Korea

ABSTRACT

Cirsium rhinoceros (H.Lév. & Vaniot) Nakai has been used a traditional medicine. Complete chloroplast genome of *C. rhinoceros* is 152,576 bp long and has four subregions: 87,262 bp of large single copy (LSC) and 21,486 bp of small single copy (SSC) regions that are separated by 18,742 bp of inverted repeat (IR) regions including 133 genes (88 protein-coding genes, 8 rRNAs, and 37 tRNAs). The overall GC content of this chloroplast genome is 37.7% and in the LSC, SSC, and IR regions are 36.0%, 31.4%, and 43.8%, respectively. Phylogenetic trees show that *Cirsium* species are clustered along with their distribution.

Cirsium rhinoceros (H.Lév. & Vaniot) Nakai, which is Korean endemic species distributed in Jeju island, Republic of Korea, has been utilized as traditional medicine. It contains useful compounds, such as polyacetylene, three flavonoids, and norisoprenoids (Lee et al. 1994; Chung et al. 2002; Soon-Ho et al. 2003). Five *Cirsium* chloroplast genomes have been sequenced including two Korean species, which have also been used as traditional medicine. The rest three are from North America and one from Europe (Dann et al. 2017), which is a good resource for understanding the phylogenetic relationship based on geographical distribution. Complete chloroplast genomes have been used for understanding the phylogenetic position of endemic species, such as genus *Pseudostellaria longipedicellata* and *Pseudostellaria palibiniana* (Kim et al. 2018; Kim and Park 2019), *Salix koriyanagi* (Kim et al. 2019), a new variety candidate of *Potentilla* (Heo et al. 2019), and etc.

To understand the phylogenetic position of *C. rhinoceros*, its total DNA isolated from Jeju island, Republic of Korea (38.0960 N 127.4762 E) was extracted from fresh leaves by using a DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany). The voucher was deposited in InfoBoss Cyber Herbarium (IN; IBS-00010). The genome was sequenced using HiSeqX at Macrogen Inc., Korea, and *de novo* assembly and confirmation were performed by Velvet 1.2.10 (Zerbino and Birney

2008), SOAPGapCloser 1.12 (Zhao et al. 2011), BWA 0.7.17 (Li 2013), and SAMtools 1.9 (Li et al. 2009). Geneious R11 11.0.5 (Biomatters Ltd., Auckland, New Zealand) was used for annotation based on *Cirsium vulgare* chloroplast genome (NC_036967; Dann et al. 2017).

The chloroplast genome of *C. rhinoceros* (Genbank accession is MK922360) is 152,576 bp long (GC ratio is 37.7%) and has four subregions: 87,262 bp of large single copy (LSC; 36.0%) and 21,486 bp of small single copy (SSC; 31.4%) regions are separated by 18,742 bp of inverted repeat (IR; 43.8%). It contains 133 genes (88 protein-coding genes, 8 rRNAs, and 37 tRNAs); 16 genes (six protein-coding genes, four rRNAs, and six tRNAs) are duplicated in IR regions.

Six *Cirsium* including *C. rhinoceros,* two *Cynara*, one *Centaurea*, and one *Saussurea* chloroplast genomes were used for constructing neighbor joining (bootstrap repeat is 10,000) and maximum likelihood (bootstrap repeat is 1,000) phylogenic trees using MEGA X (Kumar et al. 2018) after aligning whole chloroplast genomes using MAFFT 7.388 (Katoh and Standley 2013). Phylogenetic trees show that *C. rhinoceros* is clustered with two varieties of *Cirsium japonicum* distributed in the Korean peninsula and two *Cirsium* species in North America, and *Cirsium eriophorum* in Europe are separately clustered with three Korean *Cirsium* (Figure 1). It presents that *Cirsium* species have been evolved separately
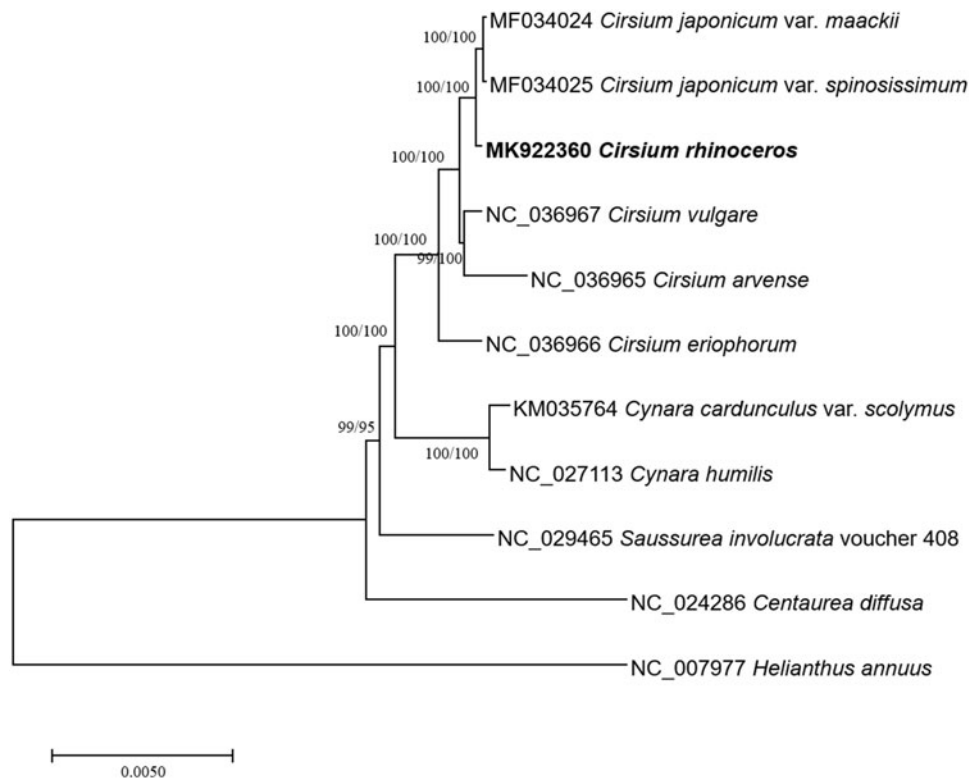

Figure 1. Neighbor joining (bootstrap repeat is 10,000) and maximum likelihood (bootstrap repeat is 1000) phylogenetic trees of six *Cirsium* including *C. rhinoceros,* two *Cynara*, one *Centaurea*, and one *Saussurea* from Asteraceae: *Cirsium rhinoceros* (MK922360 in this study), *Cirsium arvense* (NC_036965), *Cirsium eriophorum* (NC_036966), Cirsium vulgare (NC_036967), *Cirsium japonicum* var. *maackii* (MF034024), *Cirsium japonicum* var. *spinosissimum* (MF034025), *Cynara cardunculus* var. *scolymus* (KM035764), *Cynara humilis* (NC_027113), *Centaurea diffusa* (NC_024286), and *Saussurea involucrata* (NC_029465). Phylogenetic tree was drawn based on neighbor joining tree. The numbers above branches indicate bootstrap support values of maximum likelihood and neighbor joining phylogenetic tree, respectively.

along with different geographical distribution. Many Korean *Cirsium* species will give chances to understand *Cirsium* phylogenetic relationship together with available *Cirsium* chloroplast genomes: whether *Cirsium* species in Korean peninsula, as refugia in glacial epoch, have been evolved divergently in comparison to other continents or not.

Disclosure statement

The authors declare that they have no competing interests.

Funding

ORCID

Yongsung Kim 🔟 http://orcid.org/0000-0002-5349-9226
Jongsun Park 🔟 http://orcid.org/0000-0003-0786-4701

References

Chung A-K, Kwon H-C, Choi S-Z, Min Y-D, Lee S-O, Lee W-B, Yang M-C, Lee K-H, Nam J-H, Kwak J-H. 2002. Norisoprenoids from *Cirsium rhinoceros*. Kor J Pharmac. 33:81–84.

Dann M, Bellot S, Schepella S, Schaefer H, Tellier A. 2017. Mutation rates in seeds and seed-banking influence substitution rates across the angiosperm phylogeny. bioRxiv. 156398. Available from https://www.biorxiv.org/content/early/2017/06/27/156398

Heo K-I, Park J, Kim Y. 2019. The complete chloroplast genome of new variety candidate in Korea, *Potentilla freyniana* var. *chejuensis* (Rosoideae). Mitochondrial DNA Part B. 4:1354–1356.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30:772–780.

Kim Y, Heo K-I, Lee S, Park J. 2018. Complete chloroplast genome sequence of the *Pseudostellaria longipedicellata* S. Lee, K. Heo & SC Kim (Caryophyllaceae). Mitochondrial DNA Part B. 3:1296–1297.

Kim J, Kim Y, Park J. 2019. Complete chloroplast genome sequence of the *Salix koriyanagi* Kimura ex Goerz (Salicaceae). Mitochondrial DNA Part B. 4:549–550.

Kim Y, Park J. 2019. Complete chloroplast genome sequence of the *Pseudostellaria okamotoi* Ohwi (Caryophyllaceae). Mitochondrial DNA Part B. 4:174–175.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Mol Biol Evol. 35:1547–1549.

Lee HB, Kwak JH, Zee OP, Yoo SJ. 1994. Flavonoids from *Cirsium rhinoceros*. Arch Pharmacal Res. 17:273–277.

Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Genomics. arXiv:13033997.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The sequence alignment/map format and SAMtools. Bioinformatics. 25:2078–2079.

Soon-Ho Y, Jung KH, Ik-Soo L. 2003. A polyacetylene and flavonoids from *Cirsium rhinoceros*. Arch Pharmacal Res. 26:128.

Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Gen Res. 18:821–829.

Zhao Q-Y, Wang Y, Kong Y-M, Luo D, Li X, Hao P. 2011. Optimizing *de novo* transcriptome assembly from short-read RNA-Seq data: a comparative study. BMC Bioinform. 12:S2.