

The sequencing and characterization of chloroplast genome of *Geranium sibiricum* Linne

Sun Huizhen and Xinyuan Xue

Center for Ecological Research, Northeast Forestry University, Harbin, P.R. China

ABSTRACT

The complete chloroplast genome of *Geranium sibiricum* Linne. was sequenced, assembled and annotated. It is a circular form of 150,656 bp in length, which was separated into four distinct regions, a large single copy (LSC) of 73,862 bp, a small single copy region (SSC) of 52,666 bp, two inverted repeats (IR) of 12,064 bp. A total of 124 genes were predicted, of which, 87 encode proteins, 4 rRNA, 33 tRNA. The evolutionary history was inferred using Maximum Likelihood method, and the result indicates that *G. sibiricum* was grouped within Geraniaceae, and comprised a clade with *Geranium palmatum* under 100% Bootstrap value.

ARTICLE HISTORY

Received 6 March 2021
Accepted 21 June 2021

KEYWORDS

Geranium sibiricum;
chloroplast genome;
phylogenetic analysis

Geranium sibiricum L., belonging to Geraniaceae, is a widely distributed herb in northern China, Korea, Japan, and some European countries (Shim et al. 2009; Wu et al. 2010). This herb is used as food, or condiment for cooking meat in some countries (Wu et al. 2010). For containing bioactive constituents, such as geraniin, corilagin, and gallic acid, it is also used as medicines for healing bacteria, intestinal inflammation, dermatitis, diarrhea, cancer and for hair growth-promoting (Wang et al. 2015; Shim and Lim 2008, 2009; Shim et al. 2009; Wu et al. 2010; Boisvert et al. 2017). In previous studies, although many literatures on chemical composition and their functions have been documented, there were hardly any its genetic researches. In this study, we report the complete chloroplast (cp) genome of *G. sibiricum*.

Samples were collected from Qilian mountains (36°35'12"N, 101°49'11"E) in Qinghai province. Voucher specimen (LQ2020060506) was deposited in the Herbarium, Center for Ecological Research, Northeast Forestry University. A sample's total genomic DNA was extracted from about 100 mg fresh leaves using a modified CTAB method (Murray and Thompson 1980). An average length of 350 bp Paired-end Libraries with NexteraXT DNA Library Preparation Kit (Illumina, San Diego, CA) was constructed and sequenced on Illumina Novaseq 6000 platform (Shenzhen Huitong biotechnology Co. Ltd). Raw sequence reads were filtered using NGS QC Toolkit (Patel and Jain 2012) and 5.86 Gb clean data was

de novo assembled by SPAdes v.3.9.0 soft-ware (Bankevich et al. 2012) with the chloroplast genome of *Geranium incanum* Burm. f. (Accession number: KT760575.1) as reference. The assembled complete cp genome was annotated via PGA (Qu et al. 2019).

The complete cp genome of *G. sibiricum* (GenBank accession no. MW266075.1) is quadripartite form of 150,656 bp in length, and composed of a large single copy region (LSC, 73,862 bp), a small single copy region (SSC, 52,666 bp), two inverted repeats (IR, 12,064 bp). GC content of the genome is 38%. A total of 124 genes were predicted on this cp genome, of which, 87 encode proteins, 4 rRNA, 33 tRNA.

Phylogenetic analysis was performed based on complete cp genomes of *G. sibiricum* and other 18 related species reported in Geraniaceae, three species in Malvaceae as out-group. The sequences were aligned using HomBlocks (Bi et al. 2018). The evolutionary history was inferred with Maximum Likelihood (ML) method using MEGA X under GTR + G + I model, and partial deletion of gaps/missing data (Kumar et al. 2018). Bootstrap (BS) values were calculated from 1000 replicate analysis (Figure 1). As expected, *G. sibiricum* was grouped within Geraniaceae, and comprised a clade with *Geranium palmatum* under 100% BS value. The complete cp genome of *G. sibiricum* will be helpful for further studies on molecular biology, population genetics, taxonomy or resources protection.

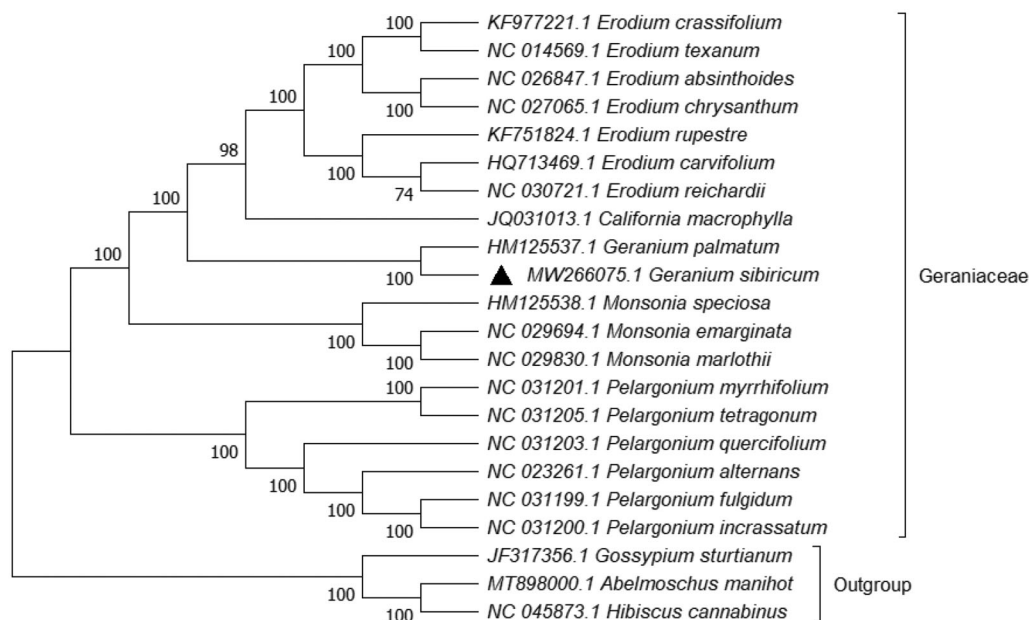


Figure 1. ML phylogenetic tree based on 22 species chloroplast genomes was constructed using MEGA X. Numbers on each node are bootstrap from 1000 replicates.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Fundamental Research Funds for the Central Universities [Grant No. 2572019CP14 and No. 2572018BA05].

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/nucleotide/MW266075.1>) under the accession no. MW266075.1. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA705079, SRR13789010, and SAMN18062704, respectively.

References

- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19(5):455–477.
- Bi GQ, Mao YX, Xing QK, Cao M. 2018. HomBlocks: a multiple-alignment construction pipeline for organelle phylogenomics based on locally collinear block searching. *Genomics.* 110(1):18–22.
- Boisvert WA, Yu M, Choi Y, Jeong GH, Zhang YL, Cho S, Choi C, Lee S, Lee BH. 2017. Hair growth-promoting effect of *Geranium sibiricum* extract in human dermal papilla cells and C57BL/6 mice. *BMC Complement Altern Med.* 17(1):109.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol.* 35(6):1547–1549.
- Murray MG, Thompson WF. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8(19):4321–4326.
- Patel RK, Jain M. 2012. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PLOS One.* 7(2):e30619.
- Qu XJ, Moore MJ, Li DZ, Yi TS. 2019. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. *Plant Methods.* 15:50.
- Shim JU, Lim K-T. 2008. Anti-oxidative and anti-proliferative character of glycoprotein isolated from *Geranium Sibiricum* Linne in Chang liver cells. *Environ Toxicol Pharmacol.* 26(3):320–324.
- Shim JU, Lim KT. 2009. Antioxidative activity of glycoprotein isolated from *Geranium sibiricum* Linne. *Nat Prod Res.* 23(4):375–387.
- Shim JU, Oh P-S, Lim K-T. 2009. Anti-inflammatory activity of ethanol extract from *Geranium sibiricum* Linne. *J Ethnopharmacol.* 126(1): 90–95.
- Wang P, Peng X, Wei ZF, Wei FY, Wang W, Ma WD, Yao LP, Fu YJ, Zu YG. 2015. Geraniin exerts cytoprotective effect against cellular oxidative stress by upregulation of Nrf2-mediated antioxidant enzyme expression via PI3K/AKT and ERK1/2 pathway. *Biochim Biophys Acta.* 1850(9): 1751–1761.
- Wu N, Zu YG, Fu YJ, Kong Y, Zhao JT, Li XJ, Li J, Wink M, Efferth T. 2010. Antioxidant activities and xanthine oxidase inhibitory effects of extracts and main polyphenolic compounds obtained from *Geranium sibiricum* L. *J Agric Food Chem.* 58(8):4737–4743.