The complete chloroplast genome of *Lilium nepalense* (Liliaceae)

Hongzhi Wu, Weiwei Bai, Shuilian He, Yan Zhao and Jingzhi Wu

Yunnan Agricultural University, Kunming, China

**ABSTRACT**

*Lilium nepalense* is a useful plant species not only for its showy flowers but also has high medicinal value. In this study, the whole chloroplast genome of *L. nepalense* was sequenced for the first time. The genome size of *L. nepalense*, was 152,956 bp, with typical tetragonal structure: one large single copy (82,573 bp), one small single copy (17,527 bp), and a pair of inverted repeat regions (IRs, 26,428 bp). The overall GC content was 37.0%. The complete genome contained 131 genes, including 85 protein-coding genes, 38 tRNA genes, and 8 rRNA genes. Phylogenetic analysis showed that *L. nepalense* was a close relationship between *L. leucanthum* and *L. henryi*. Phylogenetic analysis placed *L. nepalense* under the family Liliaceae.

**Lilium nepalense** belongs to the perennial bulbous flower of Sect of Liliaceae. Its flowers are light yellow green, with green outside and purple spots inside. *Lilium nepalense* has a light fragrance, large number of flowers per plant, strong disease resistance, and adaptability, so with high ornamental value. It is one of the most promising wildflower resources in Yunnan. At present, there are no green flower varieties in lily cultivars (Guo et al. 2019). Therefore, *L. nepalense* is one of the potential parents to cultivate new varieties of green lily.

It is widely used in the Ayurvedic medicinal system, mostly as a tonic, and has been describes as a stomachic, stimulant, and aromatic drug. The beautiful flowers are widely used for ornamental purposes. Due to the collection of large amounts of these plants, especially of the underground organs, with subsequent eradication of the natural populations, the natural habitats are rapidly decreasing (Wawrosch et al. 2001). At present, the fast propagation line of tissue culture of *L. nepalense* has been established successfully. The asexual propagation of *L. nepalense* has been carried out by tissue culture, which has enlarged the population on a large scale, not only fixed the excellent characters but also shortened the breeding cycle (Li et al. 2017). It provides the breeding technology guarantee for the protection of germplasm resources.

In this study, we determined the chloroplast genome sequence of *L. nepalense*, and discussed the genetic relationship among various species in Liliaceae.

The complete genomic DNA was extracted from fresh plant leaves and *L. rosthornii* was collected from Yuangyang County of Honghe prefecture in Yunnan province. Specimens were kept in Herbarium of Yunnan Agricultural University under the collection (No. 2020WHZ003). Total genomic DNA was isolated from fresh leaves using a DNeasy Plant Mini Kit (Qiagen, Valencia, CA) according to the manufacturer’s instructions to construction chloroplast DNA libraries. Sequencing was carried out on an Illumina NovaSeq platform. The output was a 5 Gb raw data of 150 bp paired-end reads, further trimmed and assembled using SPAdes (Bankevich et al. 2012). Resultant clean reads were assembled using GetOrganelle pipeline (https://github.com/Kinggerm/GetOrganelle). Annotations of chloroplast genome were conducted by the software Geneious (Kearse et al. 2012) and checked by comparison against the *L. leucanthum* complete chloroplast genome (GenBank accession number: KY748299).

The results of whole-genome sequencing showed that the size of chloroplast genome of *L. nepalense* (Genbank accession number: MW136391) was 152,956 bp with a typical tetragonal structure: one large single copy (LSC, 82,573 bp), one small single copy (SSC, 17,527 bp), and two reverse repeats (IRs, 26,428 bp). The total GC content was 37.0%, the GC content of LSC was 34.8%, and the GC content of SSC was 30.6%. A total of 131 genes were detected, including 85 protein-coding genes, 38 tRNA genes, and 8 rRNA genes. Nineteen gene are partially or completely duplicated, including seven PCG (rpl2; rpl23; ycf2; ndhB; rps12; rps7; ycf1), eight tRNA (trnL-GAU, trnA-UGC, trnL-CAC, trnL-CAU, trnR-ACG, trnV-GAC, trnN-GUU, trnH-GUG), and all four rRNA (4.5S, 5S, 16S, and 23S rRNA). All the rRNA genes in the genome sequence were located in the repeat region.

Based on the chloroplast genome, 24 *Lilium* species and four outgroups were used to construct the phylogenetic tree of *Lilium*. According to the general time reversible substitution model in phyml 3.0, the maximum likelihood (ML) method was used to infer the phylogenomic relationships (Larkin et al. 2007; Guindon et al. 2010). The ML tree was conducted with 1000 bootstrap replicates using FastTree software (Liu et al. 2011). According to the result of the analysis, *L. nepalense* belongs to family Liliaceae and is associated with other *Lilium* species. The phylogenetic tree showed...
that 24 Lilium species were grouped into two branches, and *L. nepalense* was related to *L. leucanthum* and *L. henryi*. It has important medicinal and ornamental value and can provide a reference for other sect. In the meantime, it provides essential data for further study on the accurately identifying species, taxonomy, and phylogenetic resolution and evolution for the genus Lilium, and the available genome information could provide insight into conservation and exploitation efforts for this endangered ornamental and medicinal species (Figure 1).

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This work was founded by Yunnan Provincial Key R&D Program [Grant no. 2018BB010], and the Academician Expert Workstation from the Yunnan Provincial Science and Technology Department [Grant no. 2018IC098], and the project of innovation and utilization for high mountain flowers by Kunming Science and Technology Bureau [Grant no. 2019-1-N-2531800003433].

**Data availability statement**

The data that support the findings of this study are openly available in GenBank of NCBI at https://www.ncbi.nlm.nih.gov/, reference number MW136391.

**References**

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pyshelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 19(5):455–477.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 59(3):307–321.

Guo PH, Wu JZ, Wang YY, Zhen YY, Wu HZ. 2019. Compatibility and overcoming of distant hybridization of *Lilium nepalense* and hybridization obstacles. J South Agric. 50(9):1903–1912.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 28(12):1647–1649.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilms A, Lopez R, et al. 2007. Clustal W and clustal X version 2.0. Bioinformatics. 23(21):2947–2948.

Li D, Luo YR, Han GW, Zhang X, Wu YY, He CZ. 2017. Studies on poly-ploidy induct-ion of wild *Lilium nepalense* D. Don. J Yunnan Univer. 32(4):678–684.

Liu K, Randal LC, Tandy W, Rongling W. 2011. RAxML and FastTree: comparing two methods for large-scale maximum likelihood phylogeny estimation. PLOS One. 6(11):e27731.

Wawrosch C, Malla PR, Kopp B. 2001. Clonal propagation of *Lilium nepalense* D.Don, a threatened medicinal plant of Nepal. Plant Cell Rep. 20(4):285–288.