**ABSTRACT**

*Agapanthus praecox* is a monocotyledonous, herbaceous, and perennial plant, which has been used as an ornamental and medicinal plant. Here, we assembled and characterized the complete chloroplast (cp) genome of *A. praecox* by de novo high throughput sequencing. The results revealed that the cp genome of *A. praecox* was 157,038 bp in total length, including a large single-copy (LSC) region of 85,195 bp, a small single-copy (SSC) region of 18,113 bp, and two invert repeats (IR) regions of 26,865 bp. The total plastid genome of *A. praecox* included 132 genes comprising 86 protein-coding genes, 38 tRNA genes, and eight rRNA genes. The phylogenetic analysis was conducted based on the complete cp genomes of 17 species and it indicated that *A. praecox* is closely related to *A. coddii* in Agapanthaceae family.

---

*Agapanthus praecox* is a monocotyledonous, herbaceous, and perennial plant (Zhang et al. 2010). This species is well-known for its aesthetic qualities which include cut flowers, potted plants, floral border plants in the landscape (Zhang et al. 2010). Besides, *A. praecox* contains medicinal compounds such as saponins, sapogenins, and phytoecdysteroids (Mori et al. 2014). APG I classification system (1998) placed the genus *Agapanthus* in the family of Agapanthaceae, then Agapanthaceae family has been subjected to Alliaceae family in APG II (2003) and APG III (2009) classification system, respectively. To date, APG IV (2016) supported APG III that Agapanthaceae was included in the expanded Amaryllidaceae family. However, there has been considerable debate over the systematic position of the genus *Agapanthus*. In the past few decades, the genus *Agapanthus* has been placed in Liliaceae family, Amaryllidaceae family, and Alliaceae family in different classification systems of angiosperm (Zhang et al. 2010). The chloroplasts (cp) genome has a maternal inheritance and conserved structure, that has been used to examine the developmental and phylogenetic relationships of plants (Wang et al. 2018). Therefore, here we assembled the complete cp genome of *A. praecox* based on Illumina pair-end sequencing data and analyzed the characteristic of its cp genome. The results of this study will provide more information for the phylogenetic reconstruction of the genus *Agapanthus*.

*A. praecox* leaves specimens were sampled from Henan Province, China (Xinyang Agriculture and Forestry University: 114° 13’ E, 32° 17’ N) and instantly preserved in liquid nitrogen. Later these specimens (YJH01) were stored in the −80°C refrigerator of Horticultural Plant Biotechnology Laboratory, Xinyang Agriculture and Forestry University. After the DNA extraction from leaf tissues, its quantification was validated by using 1% (w/v) agarose gel electrophoresis and NanoDrop spectrophotometer 2000. Qualified DNA was used for library construction and sequencing on the Illumina High-throughput sequencing platform (HiSeq2500). Approximately 12.2 GB of raw data were generated with 150 bp paired-end read lengths. The data were filtered by the script in the NOVOPlasty (Dierckxsens et al. 2016). The complete plastid genome of *Agapanthus coddii* (GeneBank accession: KX790363) was chosen as a reference. The plastid genome was assembled by GetOrganelle, it can get the plastid-like reads, and the reads were viewed and edited by Bandage (Wick et al. 2015). The cp genome annotation was assembled based on the comparison by Geneious v 11.1.5 (Biomatters Ltd, Auckland, New Zealand) (Kearse et al. 2012).

The complete cp genome sequence of *A. praecox* has been deposited to GenBank under the accession number: MW829770. Raw reads were deposited in the GenBank Sequence Read Archive (SRA: SRX10335872). The complete cp genome of *A. praecox* was a circular shape of 157,038 bp in length, with 41.97% GC content. The total plastid genome consisted of four distinct regions, such as a large single-copy (LSC) region of 85,195 bp, a small single-copy (SSC) region of 18,113 bp, and two inverted repeats (IR) regions of 26,865 bp. The complete cp genome consisted of 132 genes, including 86 protein coding genes, 38 tRNA genes, and eight rRNA genes. Phylogenetic analyses including *A. praecox*, *A. coddii*, eight Amaryllidaceae species, four Alliaceae species, and three Liliaceae species, were performed using complete cp genomes. All of the genomes were downloaded from NCBI...
GenBank. The sequences were aligned by MAFFT v7.307 (Katoh and Standley 2013), and the phylogenetic tree was constructed by MEGA X (Kumar et al. 2018). The robustness of the topology was estimated using 1000 bootstrap replicates with the neighbor-joining method (Tamura et al. 2004). The phylogenetic tree revealed that *A. praecox* is closely related to *A. coddi* in Agapanthaceae family (Figure 1).

**Disclosure statement**

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

**Funding**

This study was supported by Science and Technology Innovation Team of Xinyang Agriculture and Forestry University [CXTD202002], and the Foundation of Central Laboratory of Xinyang Agriculture and Forestry University [FCL202012].

**ORCID**

Yangyang Zhang [http://orcid.org/0000-0003-2055-4185](http://orcid.org/0000-0003-2055-4185)

Jianhua Yue [http://orcid.org/0000-0001-6573-488X](http://orcid.org/0000-0001-6573-488X)

**Data availability statement**

The data that support the findings of this study are openly available at https://www.ncbi.nlm.nih.gov/. The complete chloroplast genome has been deposited in GeneBank with accession number MW829770. And the associated Bioproject, SRA, Bio-sample numbers are PRJNA714137, SRX10335872 and SAMN18290660 respectively.

**References**

APG. 1998. An ordinal classification for the families of flowering plants. Ann Mo Bot Gard. 85(4):531–553.

APG. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. Bot J Linn Soc. 141(4):399–436.

APG. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Bot J Linn Soc. 161(2):105–121.

APG. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Bot J Linn Soc. 181(1):1–20.

Dierckxsens N, Mardulyn P, Smits G. 2016. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 45(4):e18.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780.

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.

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.

Mori S, Otani M, Kobayashi H, Nakano M. 2014. Isolation and characterization of the dihydroflavonol 4-reductase gene in the monocotyledonous ornamental *Agapanthus praecox* ssp. *orientalis* (Leighton) Leighton. Sci Hortic. 166:24–30.

Tamura K, Nei M, Kumar S. 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci USA. 101(30):11030–11035.

Wang J, Li C, Yan C, Zhao X, Shan S. 2018. A comparative analysis of the complete chloroplast genome sequences of four peanut botanical varieties. PeerJ. 6:e5349.

Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of de novo genome assemblies. Bioinformatics. 31(20):3350–3352.

Zhang D, Zhuo L, Shen X. 2010. Sporogenesis and gametogenesis in *Agapanthus praecox* Willd. *orientalis* (Leighton) Leighton and their systematic implications. Plant Syst Evol. 288(1–2):1–11.