

The chloroplast genome sequence of *Commelina communis* (Commelinaceae)

Yong Cui^{a,b} and Rifan Liang^b

^aCollege of Horticulture and Landscape, Hunan Agricultural University, Changsha, China; ^bDepartment of Agriculture and Environmental Engineering, Guangxi Vocational and Technical College, Nanning, China

ABSTRACT

Commelina communis is a hyperaccumulator on heavy metal with high drought resistance and endurance. Using Illumina next-generation sequencing data, its chloroplast genome is assembled and characterized. The complete chloroplast genome is 159,664 bp in length, consisting of a pair of inverted repeats (IRs) of 26,146 bp each, an 88,854 bp large single-copy (LSC) region and an 18,518 bp small single-copy (SSC) region. It has 137 genes in total, comprising 81 unique protein-coding genes, 32 unique tRNA genes, and 7 rRNA genes. Phylogenetic analysis based on the complete chloroplast genome sequences indicates that *C. communis* is sister to *Belosynapsis ciliata* among species with available chloroplast genome sequences.

ARTICLE HISTORY

Received 20 May 2019
Accepted 22 June 2019

KEYWORDS

Commelina communis;
complete chloroplast
genome; auto-
mated assembly

Commelina communis L., an annual herb of Commelinaceae, can grow more vigorously on the heavy metal polluted field or mine (such as copper, cadmium, lead, uranium, etc.) than on the normal unpolluted soil, with larger biomass and higher growth rate (Wang et al. 2004; Zhang 2013; Lin et al. 2014; Liu et al. 2018). It can absorb the combined form of Cu²⁺ in the whole plant, while its roots can accumulate the highest concentration of Cu (Liao et al. 2003). This makes it an ideal candidate for ecological remediation of heavy metal contaminated soil or mining habitats. However, genomic resources of this species are extremely scarce. As a first step, here we aimed to assemble its complete chloroplast genome sequence.

An individual of *C. communis* was sampled in Tongan District, Xiamen, Fujian, China (118.15°E, 24.73°N), and a voucher specimen (Cui201903) was deposited at the Sun Yat-sen University Herbarium. Genomic DNA was extracted from fresh leaves using a modified CTAB method to prepare a DNA library prior to sequencing on an Illumina HiSeq X10 platform. We finally got 2 Gbp paired-end (length = 150 bp) sequence data, which were then fed into NOVOPlasty (Dierckxsens et al. 2017) to assemble the complete chloroplast genome sequence, using *rbcL* gene sequence of *C. communis* (GenBank accession No.: MK525584) as a seed with the seed-and-extend algorithm. Online software DOGMA (Wyman et al. 2004) was used to automatically annotate the chloroplast genome, coupled with manual double-check and adjustment. We then used OGDRAW (Lohse et al. 2007) to visual the gene map of the *C. communis* chloroplast genome.

The chloroplast genome of *C. communis* (GenBank accession No.: MK863371) possessed a typical quadripartite

structure with a total length of 159,664 bp: two inverted repeats (IRs) of 26,146 bp each, an 88,854 bp large single-copy (LSC) region and an 18,518 bp small single-copy (SSC) region. It had 137 genes in total, comprising 81 unique protein-coding genes (among which seven are duplicated in the IR), 32 unique tRNA genes (seven are duplicated in the IR) and 7 rRNA genes. The genes *rps19*, *rps16*, *rpl2*, and *ndhD* had GTG, ATT, ATA, and ATC, respectively, as their start codons instead of the conventional ATG.

To infer the phylogenetic position of *C. communis*, a maximum-likelihood (ML) tree was constructed with 1000 bootstrap replicates using the GTRGAMMA substitution model (Stamatakis 2014), based on the complete chloroplast sequences of 16 species from four orders (Comelinales, Zingiberales, Poales, and Arecales) of commelinids after aligning with MAFFT v7.307 (Katoh and Standley 2013). In addition, four other species from Aparagales were used as outgroups according to the APG IV (Angiosperm Phylogeny Group et al. 2016). As shown in Figure 1, *C. communis* is sister to *Belosynapsis ciliata* among these species. The complete chloroplast genome sequence of *C. communis* will provide useful genetic resources for future studies.

Acknowledgement

We thank Nicolas Dierckxsens for assistance in using NOVOPlasty.

Disclosure statement

The authors declare that there is no conflict of interest regarding the publication of this article. The authors alone are responsible for the content and writing of the paper.

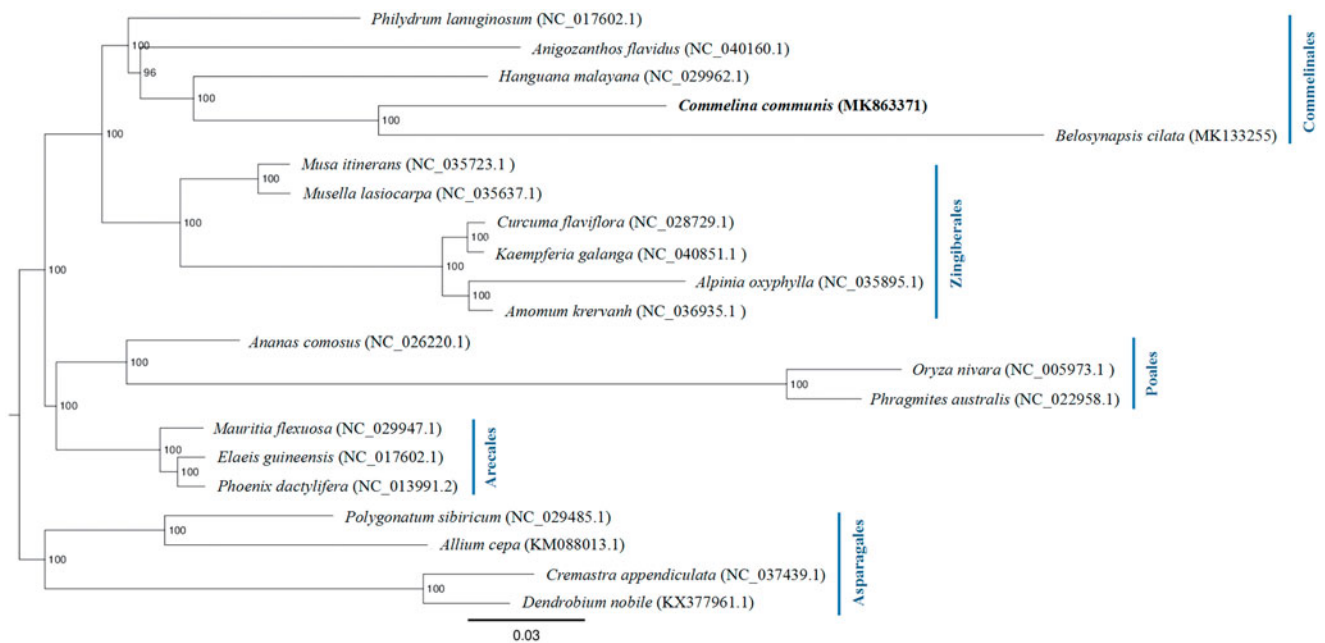


Figure 1. Maximum-likelihood tree showing the phylogenetic position of *Commelina communis* based on the complete chloroplast genome sequences. Bootstrap support values (based on 1000 replicates) are shown next to the nodes. Scale is substitutions per site.

References

- Angiosperm Phylogeny Group, Chase MW, Christenhusz MJM, Fay MF, Byng JW, Judd WS, Soltis DE, Mabberley DJ, Sennikov AN, Soltis PS, et al. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot J Linn Soc.* 181:1–20.
- Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: *de novo* assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* 45:e18.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30:772–780.
- Liao B, Deng D, Yang B, Shu W, Lin L, Lan C. 2003. [Cu tolerance and accumulation in *Commelina communis*]. *J Environ Sci.* 23:797–801.
- Lin F, Cong X, Huang J, Chen Q. 2014. [Resistance of artificial wetland plants to lead]. *Chin J Environ Eng.* 8:2329–2334.
- Liu L, Wang W, Zhang Y. 2018. [Bioconcentration and tolerance characteristics of *Commelina communis* on uranium]. *Guangdong Agr Sci.* 45: 86–92.
- Lohse M, Drechsel O, Bock R. 2007. Organellargenomedraw (ogdraw): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr Genet.* 52:267–274.
- Stamatakis A. 2014. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 30:1312–1313.
- Wang H, Shan X, Wen B, Zhang S, Wang Z. 2004. Responses of antioxidative enzymes to accumulation of copper in a copper hyperaccumulator of *Commelina communis*. *Arch Environ Contam Toxicol.* 47: 185–192.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics.* 20:3252–3255.
- Zhang X. 2013. [Cadmium tolerance and accumulation characteristics of *Commelina communis* Linn]. *Guangdong Agr Sci.* 40:167–169.