

## Complete chloroplast genome sequence of *Kosteletzkya pentacarpus*

Zhiqian Wang<sup>a</sup>, Hong Yang<sup>a</sup>, Fengjiao Zhang<sup>a</sup>, Yunlong Yin<sup>a,b</sup> and Chunsun Gu<sup>a,b,c</sup>

<sup>a</sup>Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing, China; <sup>b</sup>Jiangsu Provincial Platform for Conservation and Utilization of Agricultural Germplasm, Jiangsu Key Laboratory for the Research and Utilization of Plant Resources, Nanjing, China; <sup>c</sup>College of Forest Sciences, Nanjing Forestry University, Nanjing, China

### ABSTRACT

*Kosteletzkya pentacarpus* is a promising plant being developed as a salt-tolerant biofuel crop that also the ability to fix heavy metals. Here, high-throughput sequencing technology was used to sequence and assemble the chloroplast genome of *K. pentacarpus*. The full length of the chloroplast genome is 161,777 bp, comprising a large single-copy region of 90,019 bp, a small single-copy region of 18,978 bp, and a pair of inverted repeats of 26,390 bp. A total of 113 genes were annotated, including 79 protein-coding, 30 transfer RNA, and 4 ribosomal RNA genes. Phylogenetic analysis based on whole chloroplast genome sequences showed that *K. pentacarpus* has a close relationship with *Abelmoschus* in Malvaceae. This study increases the available genomic information on *K. pentacarpus*, and provides a basis for the rational exploitation and utilization of germplasm resources.

### ARTICLE HISTORY

Received 19 October 2021  
Accepted 19 June 2022

### KEYWORDS

*Kosteletzkya pentacarpus*;  
complete chloroplast  
genome; phylogenetic  
analysis

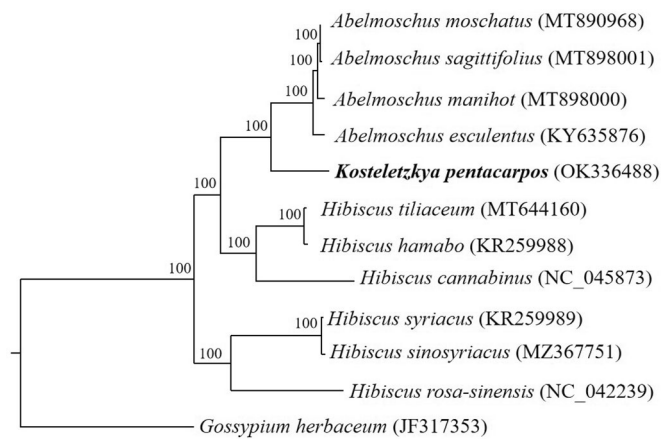
*Kosteletzkya pentacarpus* (Linnaeus) Ledebour 1842 is a perennial dicot in the family Malvaceae that is being developed as a salt-tolerant biofuel crop for salt-affected coastal land (Halchak et al. 2011; Moser et al. 2013). Its cultivation could prolong the economic viability of the land (Gallagher 1985; Voutsina 2012). This plant also has the ability to fix heavy metals through the mucilage, as has recently been demonstrated for copper, cadmium, and zinc (Lutts et al. 2016; Zhou et al. 2019; 2020). However, there are limited reports that clarify the phylogeny of this promising plant. As an effective DNA molecular marker, the chloroplast genome has been widely used in genetic and evolutionary relationship studies in plants (Freitas et al. 2018). Here, using Illumina sequencing, we determined and assembled the complete chloroplast genome sequence of *K. pentacarpus*, resulting in more sequence data, which can be used to understand the genomes and phylogenetic relationships among Malvaceae family members.

The sample of *K. pentacarpus* was collected from Nanjing Botanical Garden, Mem. Sun Yat-sen (118°49'55"E, 32°3'32"N), Nanjing, China. Total genomic DNA was extracted using a plant DNA isolation reagent (Code: D9194, TaKaRa, Dalian, China) from the fresh mature leaves of *K. pentacarpus* leaves. The specimen and DNA were deposited in Nanjing Botanical Garden, Mem. Sun Yat-sen (Zhiqian Wang, [wangzhiqian@cncbg.net](mailto:wangzhiqian@cncbg.net)) under voucher number NBG-KP-0001. A paired-end library with an insert-size of 350 bp was constructed and sequenced on an Illumina NovaSeq 6000 system (Illumina, San Diego, CA, USA). In total, approximately 8 Gb of clean data (25,236,586 reads) were obtained. The chloroplast genome was assembled using the GETORGANELLE pipeline (Jin et al. 2020) and annotated

using Geneious Prime v.2021.1.1 (<http://www.geneious.com>), with *Hibiscus cannabinus* (NC\_045873) as a reference. It was validated using CPGAVAS2 (Shi et al. 2019).

The sequence of the *K. pentacarpus* genome was deposited in GenBank (accession. OK336488). The *K. pentacarpus* chloroplast genome was determined to be 161,777 bp in length, including a large single-copy region of 90,019 bp and a small single-copy region of 18,978 bp separated by two inverted repeats of 26,390 bp. The overall GC content of the chloroplast genome is 36.9%. The genome contains 113 genes, including 79 protein-coding, 30 transfer RNA, and 4 ribosomal RNA genes. Seven protein-coding, seven transfer RNA, and four ribosomal RNA genes are duplicated in the inverted repeat region. There are 17 chloroplast genes harboring introns, with 15 containing single introns and two (*ycf3* and *clpP*) containing two introns.

The complete chloroplast genome sequences of *K. pentacarpus* and 10 other species of *Hibisceae* were used for a phylogenetic analysis, with *Gossypium herbaceum* (JF317353) as an outgroup. The chloroplast genome sequences were aligned using the MAFFT plugin (Katoh and Standley 2013) as implemented in Geneious Prime v.2021.1.1 (Biomatters Ltd., Auckland, New Zealand). The best-fitting model of nucleotide substitutions was determined using the Akaike information criterion in jModelTest v2.1.4 (Posada 2008), and the GTR + G + I substitution model was selected. Maximum-likelihood analysis was conducted using RAXML-HPC v.8.2.8 (Stamatakis 2014) with 1,000 bootstrap replicates on the CIPRES Science Gateway website (<https://www.phylo.org/>). The phylogeny revealed that *K. pentacarpus* is a sister



**Figure 1.** Phylogenetic tree, based on the chloroplast genome sequences of 12 species, showing the close relationship between *K. pentacarpus* and *Abelmoschus* L. Numbers next to the nodes indicate the bootstrap value from 1,000 replicates. The GenBank accession number for each species is shown in the parentheses after the name.

to the species of *Abelmoschus* L. (Figure 1). The chloroplast genome sequence of *K. pentacarpus* determined in this study will be useful for further analyses of molecular markers and molecular breeding.

### Author contributions

Conception and design, G.C.; Analysis and interpretation of the data, W.Z., Y.H., Z.F. and Y.Y.; Drafting of the manuscript, W.Z.; Revising the manuscript for intellectual content, Y.H., Z.F., Y.Y. and G.C.; Final approval of manuscript version to be published, W.Z., Y.H., Z.F., Y.Y. and G.C. All authors agree to be accountable for all aspects of the work.

### Ethical approval

This study did not involve endangered or protected species, and the plants were collected with special permission from Nanjing Botanical Garden, Mem. Sun Yat-sen (118°49'55"E, 32°3'32"N), Nanjing, China.

### Disclosure statement

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

### Funding

This work was supported by the Six Talent Peaks Project of Jiangsu Province [NY-042].

### 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>] under accession no. OK336488. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA767249, SRR16114382, and SAMN21895023, respectively.

### References

- Freitas A, Anuncio R, Matielo C, Stefenon VM. 2018. Chloroplast DNA: a promising source of information for plant phylogeny and traceability. *J Mol Biol Methods*. 1:1.
- Gallagher JL. 1985. Halophytic crops for cultivation at seawater salinity. *Plant Soil*. 89(1–3):323–336.
- Halchak JL, Seliskar DM, Gallagher JL. 2011. Root system architecture of *Kosteletzkya Pentacarpus* (Malvaceae) and belowground environmental influences on root and aerial growth dynamics. *Am J Bot*. 98(2): 163–174.
- Jin J-J, Yu W-B, Yang J-B, Song Y, dePamphilis CW, Yi T-S, Li D-Z. 2020. Getorganelle: a fast and versatile toolkit for accurate denovo assembly of organelle genomes. *Genome Biol*. 21(1):241.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 30(4):772–780.
- Lutts S, Qin P, Han RM. 2016. Salinity influences biosorption of heavy metals by the roots of the halophyte plant species *Kosteletzkya Pentacarpus*. *Ecol Eng*. 95:682–689.
- Moser BR, Dien BS, Seliskar DM, Gallagher JL. 2013. Seashore mallow (*Kosteletzkya Pentacarpus*) as a salt-tolerant feedstock for production of biodiesel and ethanol. *Renew Energ*. 50:833–839.
- Posada D. 2008. jModelTest: phylogenetic model averaging. *Mol Biol Evol*. 25(7):1253–1256.
- Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, Liu C. 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res*. 47(W1):W65–W73.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 30(9): 1312–1313.
- Voutsina N. 2012. Seashore mallow (*Kosteletzkya Pentacarpus*) as an ecosystem engineer: community dynamics and growth. University of Delaware, Newark, USA.
- Zhou MX, Classen B, Agneessens R, Godin B, Lutts S. 2020. Salinity improves zinc resistance in *Kosteletzkya Pentacarpus* in relation to a modification in mucilage and polysaccharides composition. *Int J Environ Res*. 14(3):323–333.
- Zhou M, Engelmann T, Lutts S. 2019. Salinity modifies heavy metals and arsenic absorption by the halophyte plant species *Kosteletzkya Pentacarpus* and pollutant leaching from a polycontaminated substrate. *Ecotox Environ Safe*. 182(OCT):109460.1–109460.10.