

## Characterization of the complete chloroplast genome of *Avena chinensis* (Poales: Poaceae)

Hangdong Wang<sup>a,b,c</sup> , Jinqing Xu<sup>a,b</sup>, Xiaolan Li<sup>a,d</sup>, En You<sup>a,d</sup>, Lei Wang<sup>a,b</sup> and Yuhu Shen<sup>a,b,c</sup>

<sup>a</sup>Key Laboratory of Adaptation and Evolution of Plateau Biota, Qinghai Provincial Key Laboratory of Crop Molecular Breeding, Laboratory for Research and Utilization of Qinghai-Tibetan Plateau Germplasm Resources, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining, China; <sup>b</sup>Innovation Academy for Seed Design, Chinese Academy of Sciences, Xining, China; <sup>c</sup>Scientific Research Station for Modern Agriculture in Wuwei Oasis, Wuwei, China; <sup>d</sup>University of Chinese Academy of Sciences, Beijing, China

### ABSTRACT

*Avena chinensis* is recognized as one of the cereals with high nutritional value in the world. In this study, the complete chloroplast (cp) genome sequence of *A. chinensis* was reported. The complete cp genome of *A. chinensis* was 135,899 bp in length with a GC content of 38.51%, including a large single copy (LSC) region of 80,117 bp, a small single copy (SSC) region of 12,576 bp, and a pair of inverted repeated regions of 21,603 bp. The *A. chinensis* cp genome encoded 128 functional genes, including 82 protein-coding genes, 38 tRNAs, and eight rRNAs. The phylogenetic analysis showed that *A. chinensis* was closely related to *Avena hybrid* and *Avena occidentalis*.

### ARTICLE HISTORY

Received 28 May 2021  
Accepted 19 September 2021

### KEYWORDS

*Avena chinensis*; chloroplast genome; *Avena*; phylogenetic tree




*Avena chinensis* (Fisch. ex Roem. & Schult.) Metzg. 1824, an annual herb of the genus *Avena* (Poaceae), is one of the most widely grown cereals in the world and a valuable resource in some countries, both for human consumption and animal feed (Fu et al. 2019). It is mainly distributed in the north, northwest and southwest of China in high latitude, high altitude, alpine arid and semi-arid areas due to its cold-loving, poor and drought-resistant characteristics (Liu et al. 2021). *A. chinensis* belonging to a tribe (Aveneae), separates from the other small-grained cereals such as wheat, barley, rye, triticale (Triticeae) and rice (Oryzae), which contains 42 chromosomes, representing three distinct sets of nuclear genomes (A, C, and D) (Marshall et al. 2013; Yan et al. 2016). The protein and fat contents of *A. chinensis* are higher than husked oats, nevertheless fiber content is lower (Givens et al. 2004; Biel et al. 2009). In the present study, the complete chloroplast (cp) genome of *A. chinensis* (GenBank accession number: MW784232) was assembled to provide genomic and genetic sources for further research.

The fresh leaves of *A. chinensis* were collected from Huangzhong (101°31' E, 36°28' N), Qinghai Province, China. Total genomic DNA of *A. chinensis* was extracted from fresh leaves using the modified CTAB method and quantified (Allen et al. 2006). The voucher specimen and extracted DNA were deposited in the Herbarium of the Northwest Institute of Plateau Biology, Chinese Academy of Sciences (Hangdong Wang, [hdwang@nwipb.cas.cn](mailto:hdwang@nwipb.cas.cn)) under the voucher number WHD2020001. Genome sequencing was performed using the

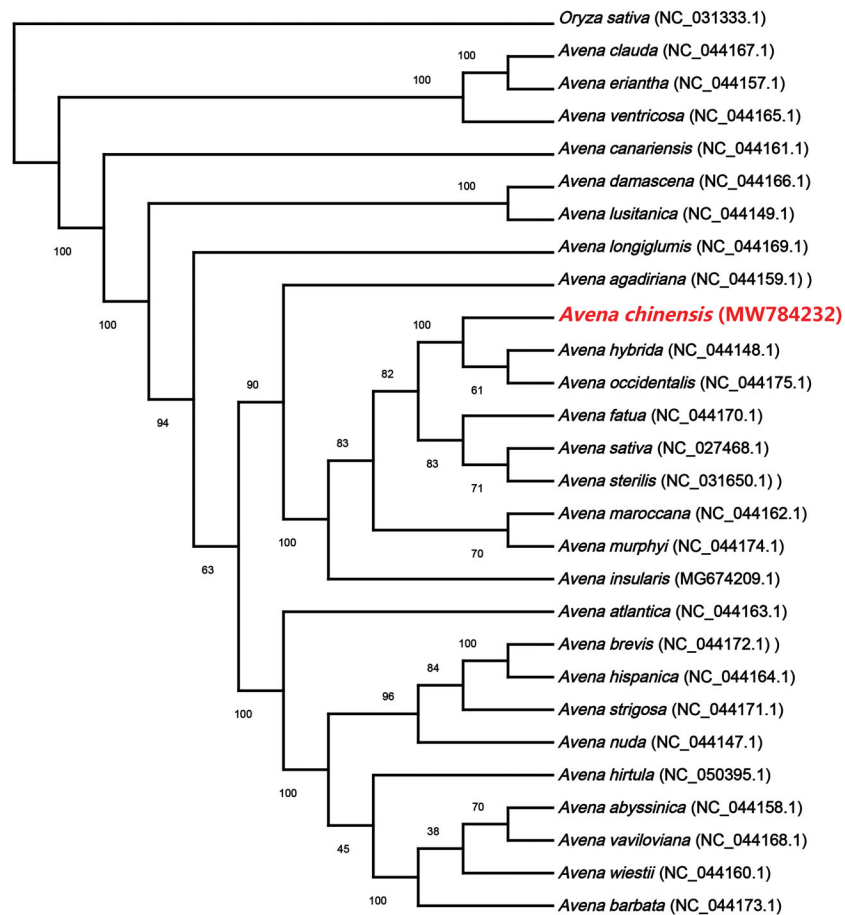
Illumina HiSeq Platform (Illumina, San Diego, CA) at Genepioneer Biotechnologies Inc. (Nanjing, China). Approximately, 26.12 million 150 bp paired-end reads were obtained, and 7.69 GB of clean data was generated after filtering. Then, the clean reads were assembled using SPAdes Version 3.10.1 (Bankevich et al. 2012), and the reference cp genome of *Avena occidentalis* (GenBank accession number: NC\_044175.1) was used for quality control after assembly. Finally, the assembled genome was annotated in CpGAVAS (Liu et al. 2012).

The complete cp genome of *A. chinensis* was 135,899 bp in length with a GC content of 38.51%, including a large single copy (LSC) region of 80,117 bp, a small single copy (SSC) region of 12,576 bp, and a pair of inverted repeated regions of 21,603 bp. The *A. chinensis* cp genome encoded 128 functional genes, including 82 protein-coding genes, 38 tRNAs, and eight rRNAs.

The maximum-likelihood phylogenetic tree (ML tree) was generated based on the complete cp genome of *A. chinensis* and 26 other species of the genus *Avena*, with *Oryza sativa* as outgroup, of which the 27 cp genomes for phylogenetic analysis were downloaded from NCBI database. The 28 complete cp genome sequences were aligned by MAFFT v7.037 (Katoh and Standley 2013). The phylogenetic tree was built using MEGA X (Kumar et al. 2018) with bootstrap set to 1000. The phylogenetic tree showed that *A. chinensis* was closely related to *Avena hybrid* and *Avena occidentalis* (Figure 1). This study was the first report on the complete cp genome of

**CONTACT** Yuhu Shen  [shenyuhu@nwipb.cas.cn](mailto:shenyuhu@nwipb.cas.cn); Lei Wang  [wanglei@nwipb.cas.cn](mailto:wanglei@nwipb.cas.cn)  Key Laboratory of Adaptation and Evolution of Plateau Biota, Qinghai Provincial Key Laboratory of Crop Molecular Breeding, Laboratory for Research and Utilization of Qinghai-Tibetan Plateau Germplasm Resources, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining, China

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Figure 1.** The ML tree based on the complete cp genome of *A. chinensis* and 26 other species of *Avena*, with *Oryza sativa* as outgroup. Numbers below or above the branches indicate the bootstrap value with 1000 replicates.

*A. chinensis* which could be useful for the phylogenetic and evolutionary studies of *Avena* and Poaceae.

and Bio-Sample numbers are PRJNA743926, SRR15044535, and SAMN20063186, respectively.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

This work was supported by the Transformation Project of Scientific and Technological Achievements of Qinghai Province under Grant No. [2019-NK-117], the Second Tibetan Plateau Scientific Expedition and Research Program (STEP) under Grant No. [2019QZKK0303], and the Construction Project for Innovation Platform of Qinghai province under Grant No. [2021-ZJ-Y05].

## ORCID

Hangdong Wang  <http://orcid.org/0000-0003-2135-5583>

## 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 the accession no. MW784232. The associated BioProject, SRA,

## References

- Allen GC, Flores-Vergara MA, Krasynanski S, Kumar S, Thompson WF. 2006. A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. *Nat Protoc.* 1(5): 2320–2325.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19(5):455–477.
- Biel W, Bobko K, Maciorowski R. 2009. Chemical composition and nutritive value of husked and naked oats grain. *J Cereal Sci.* 49(3):413–418.
- Fu YB, Li P, Biligetu B. 2019. Developing chloroplast genomic resources from 25 *Avena* species for the characterization of oat wild relative germplasm. *Plants.* 8(11):438.
- Givens DI, Davies TW, Laverick RM. 2004. Effect of variety, nitrogen fertilizer and various agronomic factors on the nutritive value of husked and naked oats grain. *Anim Feed Sci Technol.* 113(1–4):169–181.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780.
- 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.
- Liu C, Shi L, Zhu Y, Chen H, Zhang J, Lin X, Guan X. 2012. CpGAVAS, an integrated web server for the annotation, visualization, analysis, and

- GenBank submission of completely sequenced chloroplast genome sequences. *BMC Genomics*. 13(1):715.
- Liu JX, Liu RR, Jia HY, Bu T, Li N. 2021. Physiological mechanism of NaHS priming improvement of seed vigor in naked oat. *Acta Pratacult Sin*. 30(2):135–142.
- Marshall A, Cowan S, Edwards S, Griffiths I, Howarth C, Langdon T, White E. 2013. Crops that feed the world 9. Oats—a cereal crop for human and livestock feed with industrial applications. *Food Sec*. 5(1): 13–33.
- Yan HH, Bekele WA, Wight CP, Peng YY, Langdon T, Latta RG, Fu YB, Diederichsen A, Howarth CJ, Jellen EN, et al. 2016. High-density marker profiling confirms ancestral genomes of *Avena* species and identifies D-genome chromosomes of hexaploid oat. *Theor Appl Genet*. 129(11):2133–2149.