Taylor & Francis
Taylor & Francis Group

PLASTOME ANNOUNCEMENT

🔓 OPEN ACCESS ✅ Check for updates

Complete chloroplast genome of Serrated Tussock, *Nassella trichotoma* (Poaceae: Stipeae)

Aisuo Wang, Hanwen Wu 🔴 and David Gopurenko

NSW Department of Primary Industries, Wagga Wagga Agricultural Institute, PMB, Wagga Wagga, Australia

ABSTRACT

Nassella trichotoma is one of the most serious weed species in Australia. It is often confused with other *Nassella* and stipoid species, especially at the young seedling stage, adding another layer of complexity in effective weed management. We report here the complete chloroplast genome of *N. trichotoma* (137,568 bp, GenBank accession number KX792500.2) sequenced using Next Generation Sequencing technology (Illumina). The *N. trichotoma* was grouped closely with other *Nassella* species and separated from other Stipeae species in the phylogenetic tree constructed based on the complete chloroplast genome sequences. The sequence information could be used for further identification of novel DNA barcodes for correct weed identification and subsequently improve management of this invasive grass.

Nassella trichotoma (Nees) Hack. ex Arechav 1896, commonly known as Serrated Tussock, is listed as a Weed of National Significance in Australia (McLaren et al. 2002). The invasive grass causes serious damage to the environment by displacing palatable grasses from pastures, decreasing the productivity of grazing livestock, and significantly degrading biodiversity in native grasslands. Effective management of this invasive weed relies on correct weed identification, particularly at the young seedling stage. Earlier attempts at genetic diagnostics of *N. trichotoma* have been hindered by a lack of informative loci to adequately distinguish it from several other stipoid grass species (Wang et al. 2014; Wang et al. 2017). We thus sequenced the complete chloroplast genome (cp) of *N. trichotoma*. We aimed to typify the genomic content of this *Nassella* species and determine if this will facilitate the identification of novel DNA barcode regions for species identification.

Illumina sequencing technique was employed to study the complete chloroplast genome of *N. trichotoma* using a specimen collected from Wagga Wagga NSW in Australia (34°58′ 40.7″, 147° 26′ 19.5″). The specimen was taxonomically identified as *N. trichotoma* and retained at Wagga Wagga Agricultural Institute (WWAI) under voucher number ww19856 (contact person: Dr Hanwen Wu, hanwen.wu@dpi.nsw.gov.au). Total genomic DNA was extracted from fresh leaves of *N. trichotoma* using CTAB protocol (Doyle and Doyle 1990). The extracted DNA was subject to library construction (500 bp insert size) and Illumina sequencing (125 PE) at Beijing Genomics Institute (BGI, Hong Kong).

The sequencing process generated a total of 13,324,374 raw reads, which gave more than 100-fold coverage of the chloroplast genome (SRA SRP345984, BioProject PRJNA780245, BioSamples SAMN23133359). Low quality reads, adapter sequences and duplications sequences in the raw reads were trimmed with readfq v5 (https://github.com/lh3/readfq). The clean reads (13,194,420) were de novo assembled with SOAPdenovo-Trans (Xie et al. 2014) (the kmer size was optimized to 61). GapCloser (Luo et al. 2012) was applied to fill the gaps between the scaffolds using default settings. Annotation of the *N. trichotoma* chloroplast genome was performed using CPGAVAS (Liu et al. 2012), which was then followed by subsequent check with DOGMA (Wyman et al. 2004) at default settings. The predicted annotations were verified using BLAST similarity search (Altschul et al. 1990).

All annotations were manually adjusted as needed before submission to GenBank (Accession No. KX792500.2). The complete cp genome of *N. trichotoma* (137,568 bp in length) was circular in shape, having a typical quadripartite structure (LSS, SS, IRa, IRb). It consisted of 89 protein coding genes, 39 tRNA genes and eight rRNA genes, similar to other published *Nassella* cp genomes. A total of 115 SNP sites were revealed between the cp genome sequences of *N. trichotoma* and the three published *Nassella* cp genomes. Most of these SNP sites were in the genes or regions of *matK* (23), *psbA* (12), *matK-rps16* intergenic spacer (12) and *psbA-matk* intergenic spacer (10).

A total of 27 complete chloroplast sequences for the tribe Stipeae were downloaded from GenBank (accessed on 6th August 2021). These sequences were aligned with the corresponding sequences of *N. trichotoma* and an outgroup *Triticum turgidum* (tribe Triticeae) (NC_024814.1) using MAFFT
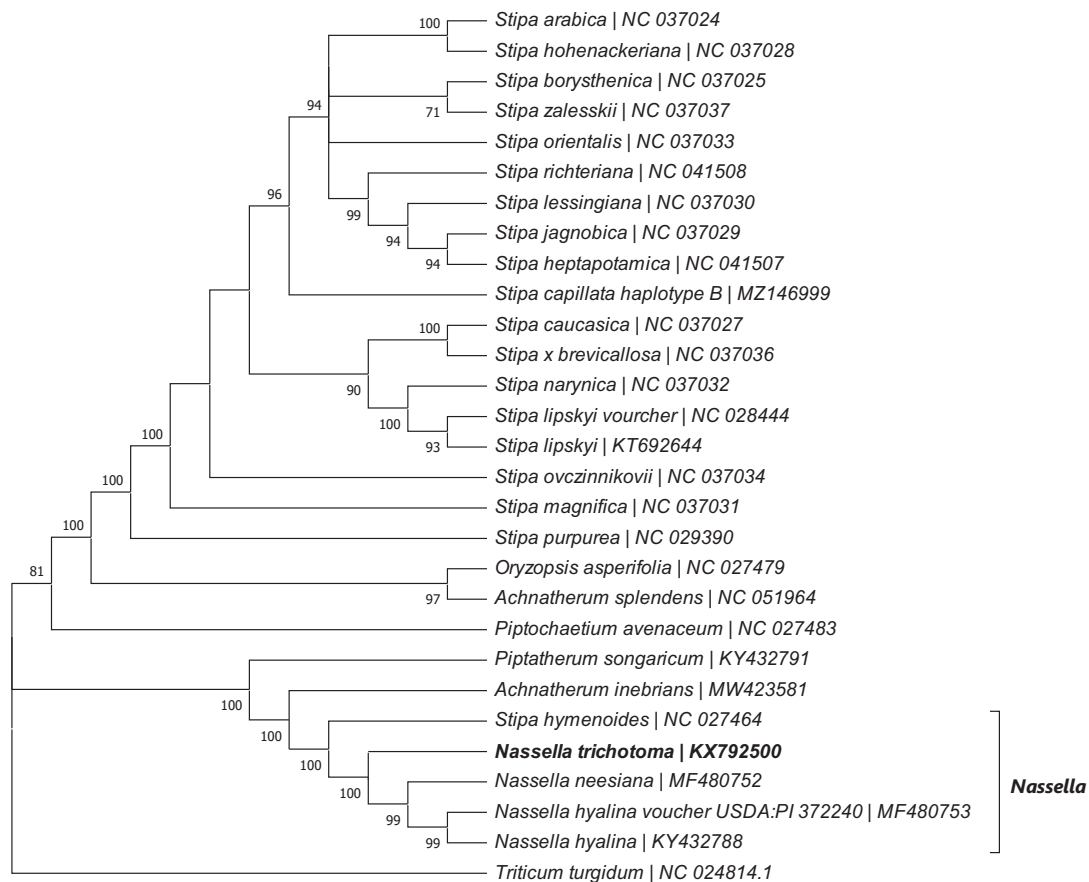
Figure 1. Maximum parsimony phylogenetic tree constructed on the complete chloroplast genomes of *N. trichotoma* and 27 other Stipeae species with *Triticum turgidum* as the outgroup.

(Katoh et al. 2005) at the default settings. The resulting alignment file was applied to construct a Maximum Parsimony (MP) tree using the Subtree-Pruning-Regrafting (SPR) algorithm (Nei and Kumar 2000) in MEGA X (Kumar et al. 2018) (Figure 1). MP clade supports were obtained by 1,000 bootstrap replications.

The resulting MP phylogenetic tree of tribe Stipeae grouped *N. trichotoma* with other *Nassella* species (100 bootstrapping support) and separated all *Nassella* species from the remaining Stipa species. One of the Stipa species, *Stipa hymenoides* (NC027464), was placed near the *Nassella* clade, which is resulted from the low mismatch numbers (639) between the cp genome of *S. hymenoides* and *N. trichotoma* (the average mismatch numbers between the cp genomes of *N. trichotoma* and the remaining Stipa species is around 1300). The close phylogenetic relationship between *S. hymenoides* and the *Nassella* requires more research to verify. In summary, the complete chloroplast genome of *N. trichotoma* provides valuable information for further genetics studies (such as DNA barcoding, evolution, and phylogeny of *Nassella* species), thereby contributing to improved understanding and management of this weed.

Acknowledgements

Authors' contribution

Ethics statement

The collection of plant specimen was conducted with the approval permit (14/8307) granted by NSW Department of Primary Industries under Section 34 of the *Noxious Weeds Act 1993*.

Disclosure statement

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

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. KX792500.2 (The associated SRA number: SRP345984, BioProject number: PRJNA780245, BioSamples number: SAMN23133359).

Funding

ORCID

Hanwen Wu 🆔 http://orcid.org/0000-0001-5157-4977

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol. 215(3):403–410.

Doyle JJ, Doyle JL. 1990. Isolation of plant DNA from fresh tissue. Focus. 12:13.

Katoh K, Kuma K, Toh H, Miyata T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. 33(2): 511–518.

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. 13:715.

Luo R, Liu B, Xie Y, Li Z, Huang W, et al. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Gigascience. 1(1):18. doi: 10.1186/2047-217X-1-18.

McLaren DA, Coventry R, Thorp J. 2002. Stipoid grasses as weeds of national significance. Plant Prot Q. 17(3):119–124.

Nei M, Kumar S. 2000. Molecular evolution and phylogenetics. New York: Oxford University Press.

Wang A, Gopurenko D, Wu H, Lepschi B. 2017. Evaluation of six candidate DNA barcode loci for identification of five important invasive grasses in eastern Australia. PLoS One. 12(4):e0175338.

Wang A, Gopurenko D, Wu H, Stanton R, Lepschi BJ. 2014. DNA barcoding for identification of exotic grass species present in eastern Australia. In: M. Baker, editor. Proceedings of the 19th Australasian weeds conference – science, community and food security: the weed challenge; Sep 1–4. Hobart (Tasmania): Tasmanian Weed Society; p. 444–447.

Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. Bioinformatics. 20(17):3252–3255.

Xie Y, Wu G, Tang J, Luo R, Patterson J, Liu S, Huang W, He G, Gu S, Li S, et al. 2014. SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. Bioinformatics. 30(12):1660–1666.