



# Characterization of the Complete Chloroplast Genome Sequences of Four *Zanthoxylum* L. Species (Sapindales: Rutaceae) from the Caribbean, Madagascar, the Mascarene Islands, and the South Pacific

 Niklas Reichelt,<sup>a,b\*</sup> Jun Wen,<sup>b</sup> Claudia Pätzold,<sup>a,c</sup> Marc S. Appelhans<sup>a,b</sup>

<sup>a</sup>Department of Systematics, Biodiversity and Evolution of Plants, Albrecht-von-Haller-Institute of Plant Sciences, University of Goettingen, Goettingen, Germany

<sup>b</sup>Department of Botany, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA

<sup>c</sup>Department of Botany and Molecular Evolution, Senckenberg Research Institute and Natural History Museum Frankfurt, Frankfurt am Main, Germany

**ABSTRACT** *Zanthoxylum* is a genus of woody plants in the Rutaceae family distributed pantropically, with some species extending to temperate regions in East Asia and North America. Here, we present the complete chloroplast genome sequences of four species, two of them critically endangered, endemic to tropical islands.

**Z***anthoxylum* L. (prickly ash, Sichuan pepper) is the second largest genus within the Rutaceae and is distributed pantropically, with some species in temperate Asia and North America (1). The genus occurs on a variety of island systems, including, e.g., the remote Hawaiian, Austral, and Juan Fernández Islands. This study provides genetic resources for the following four *Zanthoxylum* species: *Z. tragodes*, from the Caribbean and Venezuela (2); *Z. madagascariense*, endemic to Madagascar (3); *Z. pinnatum*, from the South Pacific, with the category critically endangered (CR) proposed for the Austral Island population (4); and *Z. paniculatum*, endemic to Rodrigues Island (Mauritius), with just a few specimens known today (5). Applying the International Union for Conservation of Nature (IUCN) red list criteria (6), *Z. paniculatum* should be considered CR.

Total DNA was extracted from 1 cm<sup>2</sup> herbarium leaf material from *Z. madagascariense* (collector and number, Capuron 28595-SF; country of origin, Madagascar; year of collection, 1968; herbarium, MO [herbarium codes are assigned according to the Index Herbariorum, <http://sweetgum.nybg.org/science/ih>]), *Z. paniculatum* (collector and number, C. Magdalena 001; country of origin, Mauritius [Rodrigues]; year of collection, 2007; herbarium, MO), *Z. pinnatum* (collector and number, Drake 282; country of origin, Tonga [Vava'u group]; year of collection, 1995; herbarium, US), and *Z. tragodes* (collector and number, Liogier 12644; country of origin, Dominican Republic; year of collection, 1968; herbarium, US) using the cetyltrimethylammonium bromide (CTAB) method (7). The NEBNext Ultra II DNA library prep kit for Illumina (New England Biolabs, USA) and NEBNext multiplex oligos for Illumina (dual-index primers set 1) were employed for size selection and adapter index ligation. Bait and library hybridization were performed according to the myBaits Hybridization Capture for Targeted NGS manual v.4.01 (Arbor Biosciences, USA), using a bait set designed for nuclear genes of *Zanthoxylum* and related genera (8). Plastome genome data gathered in this process were off-target hits. The captured DNA libraries were amplified with the KAPA HiFi HotStart ReadyMix (Hoffmann-La Roche, Switzerland). Target-enriched libraries were sequenced on a HiSeq 4000 instrument (Illumina, Inc., USA), producing 2 × 150-bp paired-end reads.

Trimmomatic v.0.33 (9) was used for adapter and quality trimming (PHRED33 score, >30; minimum length after trimming, 65 bp). FastUniq v.1.1 (10) was employed for

**Citation** Reichelt N, Wen J, Pätzold C, Appelhans MS. 2021. Characterization of the complete chloroplast genome sequences of four *Zanthoxylum* L. species (Sapindales: Rutaceae) from the Caribbean, Madagascar, the Mascarene Islands, and the South Pacific. *Microbiol Resour Announc* 10:e00399-21. <https://doi.org/10.1128/MRA.00399-21>.

**Editor** Antonis Rokas, Vanderbilt University

**Copyright** © 2021 Reichelt et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Niklas Reichelt, [niklas.reichelt@uni-wuerzburg.de](mailto:niklas.reichelt@uni-wuerzburg.de).

\* Present address: Niklas Reichelt, Pharmaceutical Biology, Julius-von-Sachs-Institute of Biosciences, Biocenter, University of Wuerzburg, Wuerzburg, Germany.

**Received** 21 April 2021

**Accepted** 3 May 2021

**Published** 27 May 2021

deduplication. Using Fast-Plast v.1.2.8 (11), reads were mapped to all Sapindales plastomes available in GenBank with a required minimum coverage of  $5\times$ . Annotations were drawn from the *Zanthoxylum bungeanum* (GenBank accession number [KX497031](#)) and *Phellodendron amurense* ([NC\\_035551](#)) references in Geneious v.10.0.8 using the “live annotate and predict” function, followed by manual annotation of unrecognized regions. The chloroplast genomes were visualized using OrganellarGenomeDRAW (12).

Sequencing produced between 9,939,558 and 15,323,260 reads per sample after quality trimming. In all, 25,098 to 79,647 reads mapped to the *Z. bungeanum* reference genome. This corresponds to a plastome coverage of about  $47\times$  to  $151\times$ . All plastomes show the typical structure with a large single-copy (LSC) and small single-copy (SSC) region separated by two inverted repeats (IRA and IRB). Plastome lengths range from 158,265 bp in *Z. paniculatum* to 158,722 bp in *Z. pinnatum*. The total GC contents are consistent for all four plastomes at 38.4%. In each plastome, a total of 132 genes were annotated, 40 of which are situated in the IRA and IRB regions. Genes in the two inverted repeat regions comprise all 8 rRNAs, 14 of 37 tRNAs, and 20 of 87 protein-coding genes. The gene *ycf1* within the IRA differs from its copy within the IRB, as it deeply extends into the SSC region. The protein-coding gene *ndhF* is partly situated within an inverted repeat. The protein-coding gene *rps12* is trans-spliced, and its two transcripts are found in the LSC and IR regions.

**Data availability.** The chloroplast genome sequence data were deposited in GenBank (*Z. madagascariense*, accession number [MN968551](#); *Z. paniculatum*, [MN968552](#); *Z. pinnatum*, [MN968553](#); *Z. tragodes*, [MN968554](#)). The demultiplexed raw data were deposited at the NCBI Sequence Read Archive (BioProject accession number [PRJNA623909](#)).

## ACKNOWLEDGMENTS

We thank the curators of MO for allowing us to sample their collections of *Zanthoxylum*.

This study was supported by funds from the Smithsonian Institution Fellowship Program and PROMOS (DAAD).

## REFERENCES

- Kubitzki K, Kallunki JA, Duretto M, Wilson PG. 2011. Rutaceae, p 276–356. In Kubitzki K (ed), The families and genera of vascular plants. Springer, Berlin, Germany.
- Reynel C. 2017. Flora neotropica monograph no. 117. *Zanthoxylum* (Rutaceae). New York Botanical Garden Press, New York, NY.
- Perrier de la Bâthie H. 1950. Rutacées, p 1–89. In Humbert H (ed), Flore de Madagascar et des Comores. Firmin-Didot et Cie, Paris, France. (In French.)
- Butaud JF, Meyer JY. 2004. Plans de conservation pour des plantes menacées et/ou protégées en Polynésie française. Contribution à la biodiversité de Polynésie française, vol 11. Service du Développement Rural/Développement à la Recherche, Papeete, French Polynesia. (In French.)
- Bone R. 2005. A proposal for rare plant rescue: *Zanthoxylum paniculatum* Balf. fil. (Rutaceae), endemic to Rodrigues. <https://www.dendrology.org/publications/dendrology/a-proposal-for-rare-plant-rescue-zanthoxylum-paniculatum-balf-fil-rutaceae-endemic-to-rodrigues/>.
- International Union for Conservation of Nature. 2019. The IUCN red list of threatened species version 2018-2. International Union for Conservation of Nature, Cambridge, United Kingdom. <http://www.iucnredlist.org>.
- Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* 19:11–15.
- Reichelt N, Wen J, Pätzold C, Appelhans MS. 2021. Target enrichment improves phylogenetic resolution in the genus *Zanthoxylum* (Rutaceae) and indicates both incomplete lineage sorting and hybridization events. *bioRxiv* <https://doi.org/10.1101/2021.04.12.439519>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Xu H, Luo X, Qian J, Pang X, Song J, Qian G, Chen J, Chen S. 2012. FastUniq: a fast de novo duplicates removal tool for paired short reads. *PLoS One* 7:e52249. <https://doi.org/10.1371/journal.pone.0052249>.
- McKain MR, Wilson M. 2017. Fast-Plast: rapid de novo assembly and finishing for whole chloroplast genomes. <https://github.com/mrmckain/Fast-Plast>.
- Lohse M, Drechsel O, Kahlau S, Bock R. 2013. OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualization expression data sets. *Nucleic Acids Res* 41:W575–W581. <https://doi.org/10.1093/nar/gkt289>.