

The complete chloroplast genome sequence of *Populus mexicana* (Salicaceae)

Zhe Hou

Key Laboratory of Southwest China Wildlife Resources Conservation (Ministry of Education), College of Life Science, China West Normal University, Sichuan, Nanchong, China

ABSTRACT

The complete chloroplast genome sequence of *Populus mexicana* was characterized from Illumina pair-end sequencing. The chloroplast genome of *P. mexicana* was 156,188 bp in length, containing a large single-copy region (LSC) of 86,871 bp, a small single-copy region (SSC) of 16,559 bp, and 2 inverted repeat (IR) regions of 26,377 bp. The overall GC content is 36.74%, while the corresponding values of the LSC, SSC, and IR regions are 64.8, 69.3, and 60.3, respectively. The genome contains 111 complete genes, including 72 protein-coding genes (62 protein-coding gene species), 31 tRNA genes (29 tRNA species), and 8 rRNA genes (4 rRNA species). The neighbor-joining phylogenetic analysis showed that *P. mexicana* and *Populus fremontii* clustered together as sisters to other *Populus* species.

ARTICLE HISTORY

Received 14 November 2019
Accepted 23 November 2019

KEYWORDS

Populus mexicana;
chloroplast genome;
phylogenetic analysis;
genetic information

Introduction

Populus mexicana, a rare tree, is the sole extant species of *Populus* L. sect. *Abaso* Eckenw. Although *P. mexicana* is geographically restricted to the warmer areas of southern North America (NA) today, early fossil distributions reveal that ancestral poplars of section *Abaso* were widespread across NA and as far north as Alaska (Eckenwalder 1977). *Populus mexicana* can adapt to different climates and environments owing to without anthropogenic influence and harbor a wealth of genetic variation. *Populus mexicana* harbors high ecological and economic value and have high level of intra-specific genetic diversity (Callahan et al. 2013). Therefore, *P. mexicana* is an excellent system for understanding genetic information and genome variation patterns (Neale and Antoine 2011). Moreover, we can develop conservation strategies easily when we understand the genetic information of *P. mexicana*. In the present research, we constructed the whole chloroplast genome of *P. mexicana* and understood many genome variation information about the species, which will provide beneficial help for population genetics studies of *P. mexicana*.

The fresh leaves of *P. mexicana* were collected from Mexico City (32°13'N, 110°57'W). Fresh leaves were silica-dried and taken to the laboratory until DNA extraction. The voucher specimen (PME001) was laid in the Herbarium of China West Normal University and the extracted DNA was stored in the -80 °C refrigerator of the Key Laboratory of Southwest China Wildlife Resources Conservation. We extracted total genomic DNA from 25 mg silica-gel-dried leaf

using a modified CTAB method (Doyle 1987). The Illumina HiSeq 2000 platform (Illumina, San Diego, CA, USA) was used to perform the genome sequence. We used the software MITObim 1.8 (Hahn et al. 2013) and metaSPAdes (Nurk et al. 2017) to assemble chloroplast genomes. We used *P. tremula* (GenBank: NC_027425) as a reference genome. We annotated the chloroplast genome with the software DOGMA (Wyman et al. 2004) and then corrected the results using Geneious 8.0.2 (Campos et al. 2016) and Sequin 15.50 (<http://www.ncbi.nlm.nih.gov/Sequin/>).

The complete chloroplast genome of *P. mexicana* (GenBank accession number MN733732) was 156,188 bp in length, containing a large single-copy region (LSC) of 86,871 bp, a small single-copy region (SSC) of 16,559 bp, and two inverted repeat (IR) regions of 26,377 bp. The overall GC content is 36.74%, while the corresponding values of the LSC, SSC, and IR regions are 64.8, 69.3, and 60.3%, respectively. The genome contains 111 complete genes, including 80 protein-coding genes (70 protein-coding gene species), 31 tRNA genes (29 tRNA species), and 8 rRNA genes (4 rRNA species). Most of the genes occur as a single copy, except for 22 gene species occur in double copies.

We used the complete chloroplast genome sequence of *P. mexicana* and 12 other related species of *Populus* and *Salix* interior as outgroup to construct phylogenetic tree. The 14 chloroplast genome sequences were aligned with MAFFT (Kato and Standley 2013) and then the neighbour-joining tree was constructed by MEGA 7.0 (Kumar et al. 2016). The results confirmed that *P. mexicana* was clustered with *P. fremontii* (Figure 1).

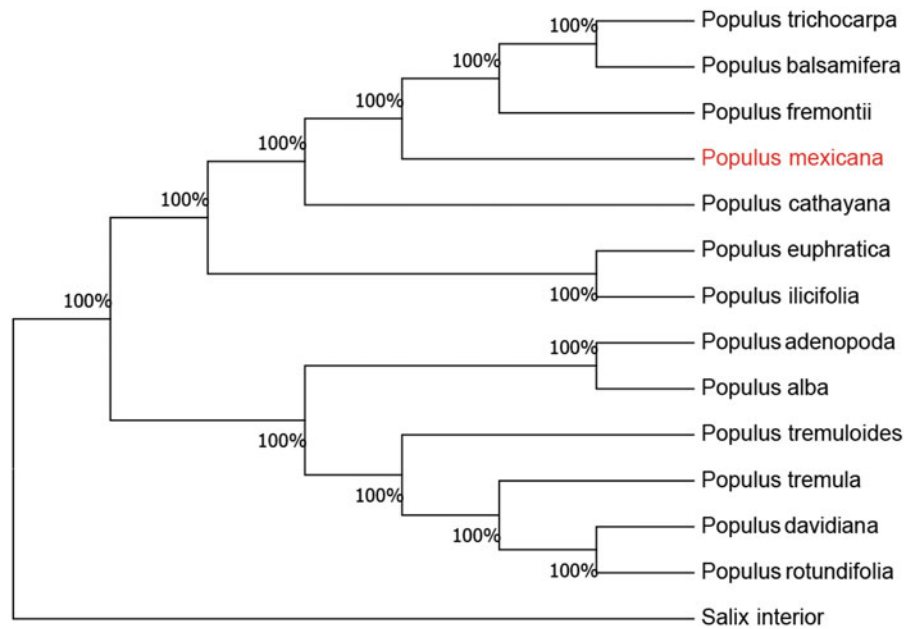


Figure 1. Neighbor-joining (NJ) analysis of *P. mexicana* and other related species based on the complete chloroplast genome sequence. Genbank accession numbers: *P. tremula* (KP861984), *P. davidiana* (KX306825), *P. yunnanensis* (KP729176), *P. euphratica* (KJ624919), *P. adenopoda* (NC032368), *P. rotundifolia* (KX425853), *P. cathayana* (KP929175), *P. balsamifera* (KJ664927), *P. ilicifolia* (NC031371), *P. trichocarpa* (EF489041), *P. fremontii* (KJ664926), *P. tremuloides* (MN561844) and *Salix interior* (NC024681).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Financial support for this research was provided by the Fundamental Research Funds of China West Normal University.

References

- Callahan CM, Rowe CA, Ryel RJ, Shaw JD, Madritch MD, Mock KE. 2013. Continental-scale assessment of genetic diversity and population structure in quaking aspen (*Populus tremuloides*). *J Biogeogr.* 40(9): 1780–1791.
- Campos F, Kluge M, Franco A, Giongo A, Valdez F, Saggi T, Brito W, Roehe P. 2016. Complete genome sequence of porcine parvovirus 2 recovered from swine sera. *Genome Announc.* 4(1):e01627–01615.
- Doyle J. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull.* 19(1):11–15.
- Eckenwalder JE. 1977. North American cottonwoods (*Populus*, Salicaceae) of sections Abaso and Aigeiros. *J Arnold Arbor.* 58(3):193–208.
- Hahn C, Bachmann L, Chevreur B. 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucleic Acids Research.* 41(13):e129–e129.
- 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, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol.* 33(7):1870.
- Neale DB, Antoine K. 2011. Forest tree genomics: growing resources and applications. *Nat Rev Genet.* 12(2):111–122.
- Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* 27(5):824–834.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics.* 20(17):3252–3255.