


The complete chloroplast genome of the threatened Napa False Indigo *Amorpha californica* var. *napensis* Jeps. 1925 (Fabaceae) from Northern California, USA

Hartnell College Genomics Group^{a*}, Ivan D. Agudelo^a, Griselda Aldaco^a, Angel Brito-Pizano^a, Kimberly G. Chavez^a, Karina G. Cortina^a, Jorge Flores^a, Alejandro Fuentes^a, Adam N. Garcia^a, Alejandro Garcia^a, Daniel Gonzalez-Martinez^a, Jennifer Hernandez Ramos^a, Jeffery R. Hughey^a , Fernando R. Katada^a, Felix A. Leon^a, Maleny P. Lopez^a, Sandra Z. Lopez^a, Aileen G. Mendoza^a, Maritta Molina^a, Asmahan Muhrram^a, Daisy Ortiz-Matias^a, Tonantzin E. Ortiz^a, Alicia Pacheco^a, Nandini Patel^a, Paz M. Ramirez^a, Jennifer L. Scaramuzzino^a, Alexandria Soto^a, Richard A. Stabler^b, Jessica M. Vidauri^a, Jose Villicana^a and James A. Yhip^a

^aDivision of Mathematics, Science, and Engineering, Hartnell College, Salinas, CA, USA; ^bCounty of Sonoma, Santa Rosa, CA, USA

ABSTRACT

Amorpha californica var. *napensis* Jeps. 1925, the Napa false indigo, is a threatened shrub endemic to northern California. Here the complete chloroplast genome of topotype material of var. *napensis* was assembled and characterized to contribute to the bioinformatics, systematics, and conservation of this variety. The chloroplast genome (GenBank accession OK274088) is 158,294 base pairs (bp) in length, encodes 130 genes including 85 protein-coding, 37 tRNA, 8 rRNA, and shows a high-level of gene synteny to other Papilionoideae. Phylogenetic analysis fully resolved var. *napensis* in a clade with *A. fruticosa* L. and *A. roemeriana* Scheele, sister to the Dalbergieae. The newly sequenced chloroplast genome shows that the genetic differences between var. *napensis* and *Amorpha californica* Nutt. var. *californica* are greater than the variation observed between var. *napensis* and many other *Amorpha* spp. sequences deposited in GenBank. These data suggest that var. *napensis* should be elevated to full species rank.

ARTICLE HISTORY

Received 4 November 2021
Accepted 9 January 2022



KEYWORDS

Amorpha; *Amorpha californica*; *Amorpha californica* var. *californica*; chloroplast genome; Papilionoideae

Amorpha californica var. *napensis* is a deciduous shrub originally described by W.L. Jepson from specimens collected from Moore Creek, Howell Mountain, Napa County, California (Jepson 1925). The variety was said to differ from *A. californica* Nutt. var. *californica* in being subglabrous (vs. minutely pubescent), lacking glands on the rachis (vs. prickle-like glands), having shorter racemes 2.5–3.18 cm in length (vs. 5–14 cm), and displaying minute teeth on the calyx (vs. long teeth) (Jepson 1925). Abrams (1944) considered var. *napensis* a form of *A. californica* var. *hispidula* (Greene) E.J. Palmer, whereas Munz (1959) treated the latter as a synonym of var. *napensis*. Wilbur (1975) accepted the varietal status of var. *napensis*, but placed var. *hispidula* in synonymy under *A. californica*. Hickman (1993) and Baldwin et al. (2012) recognized the variety, but narrowed its distribution to Napa, Marin, and Sonoma counties where it occurs in Chaparral communities less than 800 meters in elevation. Calflora (<https://www.calflora.org/>) currently lists var. *napensis* as 1B.2 (fairly threatened in California). Based on a survey of this variety from the type locality, it appears nearly extirpated due to vineyard expansion in the famed Howell Mountain American Viticultural Area. To date, the only data deposited in GenBank for var. *napensis* are three barcode sequences

determined from a single specimen (Straub and Doyle 2014). We assembled and analyzed the complete plastid genome of topotype material of var. *napensis* to contribute to the bioinformatics, systematics, and conservation efforts of this threatened variety.

The specimen of var. *napensis* analyzed in this study was collected in accordance with guidelines provided by Napa county and Hartnell College from the north end of Moore Creek, Angwin, California (38°33'55.0152"N 122°24'18.8028"W) and deposited in the herbarium at Hartnell College (<https://www.hartnell.edu/>, Jeffery R. Hughey, jhughey@hartnell.edu) under voucher number HCC 266. The DNA was extracted following the methods outlined in Hughey et al. (2019). The 150 bp PE Illumina library construction and sequencing was performed by Quick Biology (Pasadena, California, USA) and yielded 24,733,870 reads. The adapters and low quality reads were removed using the Trim Adapters and Trim Low Quality default settings with the BBDuk plugin in Geneious Prime version 2019.1.3 (Biomatters Limited, Auckland, New Zealand). The genome was assembled by mapping the reads onto the reference genome *A. fruticosa*, GenBank accession number MN709789 (Zhang, Wang, et al. 2020), using the Medium-Low Sensitivity/Fast

CONTACT Jeffery R. Hughey  jhughey@hartnell.edu  Division of Mathematics, Science, and Engineering, Hartnell College, Salinas, CA, USA

*All authors contributed equally to the analysis and writing of this paper

© 2022 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.

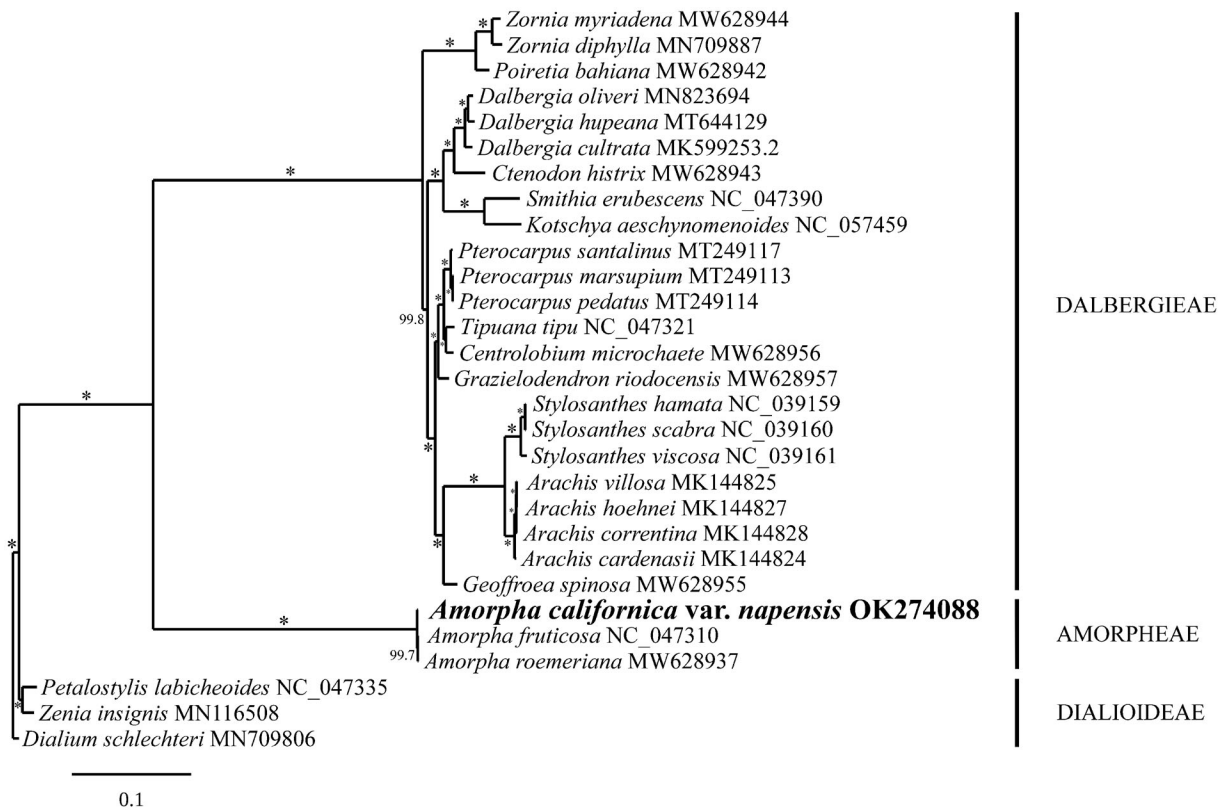


Figure 1. RaxML phylogram of the complete chloroplast genome of *Amorpha californica* var. *napensis* and related Papilionoideae. Numbers along the branches are ML bootstrap supports based on 1000 replicates (* indicates 100% bootstrap support). The legend below represents the scale for nucleotide substitutions. The following accessions with references were used for the phylogenetic analysis: *Amorpha roemeriana* MW628937, *Centrolobium microchaete* MW628956, *Ctenodon histrix* MW628943, *Geoffroea spinosa* MW628955, *Grazielodendron riodescensis* MW628957, *Poiretia bahiana* MW628942, *Zornia myriadena* MW628944 (Lee et al. 2021); *Amorpha fruticosa* NC_047310, *Dialium schlechteri* MN709806, *Petalostylis labicheoides* NC_047335, *Smithia erubescens* NC_047390, *Tipuana tipu* NC_047321, *Zornia diphylla* MN709887 (Zhang, Wang, et al. 2020); *Pterocarpus marsupium* MT249113, *Pterocarpus pedatus* MT249114, *Pterocarpus santalinus* MT249117 (Hong et al. 2020); *Stylosanthes hamata* NC_039159, *Stylosanthes scabra* NC_039160, *Stylosanthes viscosa* NC_039161 (Marques et al. 2018); *Arachis cardenasii* MK144824, *Arachis hoehnei* MK144827, *Arachis villosa* MK144825 (Wang et al. 2019); *Dalbergia hupeana* MT644129 (Hong et al. 2021); *Dalbergia oliveri* MN823694 (Zhang, Li, et al. 2020); *Kotschya aescynomenoides* NC_057459 (Oyebanji et al. 2020); *Dalbergia cultrata* MK599253 (Liu et al. 2019); *Zenia insignis* MN116508 (Lai et al. 2019); *Amorpha californica* var. *napensis* OK274088 (this study).

setting in Geneious Prime. The gaps were closed by iterative mapping using the same settings in Geneious Prime. The annotation was performed using the default settings in GeSeq (Tillich et al. 2017) and CPGAVAS2 (Shi et al. 2019), followed by adjustments according to NCBI ORFfinder, Sequin 15.5, and tRNAscan-SE 1.21 (Schattner et al. 2005). The var. *napensis* complete chloroplast nucleotide sequence was aligned to 25 other papilionoid and three outgroup taxa from the Dialioideae using the auto settings in MAFFT (Kato and Standley 2013). The ML phylogenetic analysis was executed with the TVM + F + I + G4 substitution model and 1000 ultrafast bootstrap replicates in W-IQ-TREE (Trifinopoulos et al. 2016). The tree was visualized with TreeDyn 198.3 at Phylogeny.fr (Dereeper et al. 2008).

The complete chloroplast genome of var. *napensis* is 158,294 bp in length and exhibits a standard quadripartite structure (Shinozaki et al. 1986; Wicke et al. 2011). The genome contains an LSC, SSC, and two IRs with lengths 88,110 bp, 18,580 bp and 25,802 bp, respectively. The GC content is 36.0%. Gene content and organization show a high-level of synteny to *A. fruticosa* and *A. roemeriana* (GenBank accession number MW628937) (Zhang, Wang, et al. 2020; Lee et al. 2021). The chloroplast genome of var. *napensis* is 99.86% similar in nucleotide sequence to *A. fruticosa* and 99.87% to *A.*

roemeriana. BLAST analysis of the var. *napensis* genome found an identical petN-psbM intergenic spacer sequence and nearly identical trnT-trnD intergenic spacer sequence that differed by 1 bp from var. *napensis* from Angwin, California (Straub and Doyle 2014). In comparison, var. *napensis* differed from var. *californica* from the Santa Rosa Mountains, California in its petN-psbM sequence by 5 bp (99.0%) and trnT-trnD sequence by 12 bp (99.24%). The var. *napensis* sequences were greater in genetic distance to var. *californica* than to many other species classified in *Amorpha*.

Phylogenetic analysis of representative Papilionoideae fully resolved var. *napensis* in a clade with *A. fruticosa* and *A. roemeriana* in a sister position to the Dalbergieae (Figure 1). These results are consistent with the evolutionary relationships inferred for the Amorpheae based on multigene analysis (McMahon and Hufford 2004), *matK* sequencing (Cardoso et al. 2012, 2013), and a schematic compilation based on other published works (The Legume Phylogeny Working Group 2013). Given the high degree of plastid marker sequence variation between var. *napensis* and var. *californica*, phylogenetic analysis of barcode markers or the chloroplast genome of var. *californica* from its type locality in Santa Barbara, California, are necessary to test the hypothesis that var. *napensis* should be recognized at the species level.

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

Funding

This research was supported by NSF award number 1832446 to Hartnell College.

Author contributions

All authors were equally involved in the analysis and interpretation of the data; the drafting of the paper; revising it critically for intellectual content; and the final approval of the version to be published. The corresponding author Jeffery R. Hughey and coauthor Richard A. Stabler jointly conceived and designed the project. All authors agreed to be accountable for all aspects of this work.

Data availability statement

The genome sequence data that support the findings of this study are openly available in GenBank at (<https://www.ncbi.nlm.nih.gov/>) under the accession number OK274088. The associated BioProject, SRA, and BioSample numbers are PRJNA765780, SRR16037173, and SAMN21583931 respectively.

ORCID

Jeffery R. Hughey  <http://orcid.org/0000-0003-4053-9150>

References

- Abrams L. 1944. Illustrated flora of the Pacific states: Washington, Oregon, and California. Vol. 2. Polygonaceae to Krameriaceae. Stanford (CA): Stanford University Press.
- Baldwin BG, Goldman DH, Keil DJ, Patterson R, Rosatti TJ, Wilken DH, editors. 2012. The Jepson manual: vascular plants of California. 2nd ed. Berkeley (CA): University of California Press.
- Cardoso D, de Queiroz LP, Pennington RT, de Lima HC, Fonty E, Wojciechowski MF, Lavin M. 2012. Revisiting the phylogeny of *Papilionoid legumes*: new insights from comprehensively sampled early-branching lineages. *Am J Bot.* 99(12):1991–2013.
- Cardoso D, Pennington RT, de Queiroz LP, Boatwright JS, Van Wyk BE, Wojciechowski MF, Lavin M. 2013. Reconstructing the deep-branching relationships of the *Papilionoid legumes*. *S Afr J Bot.* 89:58–75.
- Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, et al. 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* 36(Web Server issue):W465–W469.
- Hickman JC. 1993. The Jepson manual: higher plants of California. Berkeley (CA): University of California Press.
- Hong Z, Peng D, He W, Zhang N, Yang Z, Tembrock LR, Wu Z, Liao X, Xu D. 2021. Comparative analyses of 35 complete chloroplast genomes from the genus *Dalbergia* (Fabaceae) and the identification of DNA barcodes for tracking illegal logging and counterfeit rosewood. *Int J Mol Sci.* 21(11):3758 (1–18).
- Hong Z, Wu Z, Zhao K, Yang Z, Zhang N, Guo J, Tembrock LR, Xu D. 2020. Comparative analyses of five complete chloroplast genomes from the genus *Pterocarpus* (Fabaceae). *IJMS.* 21(11):3758.
- Hughey JR, Maggs CA, Mineur F, Jarvis C, Miller KA, Shabaka SH, Gabrielson PW. 2019. Genetic analysis of the Linnaean *Ulva lactuca* (Ulvales, Chlorophyta) holotype and related type specimens reveals name misapplications, unexpected origins, and new synonymies. *J Phycol.* 55(3):503–508.
- Jepson WL. 1925. A manual of the flowering plants of California. Berkeley (CA): Associated students store, University of California, Berkeley.
- Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780.
- Lai Q, Tu T, Zhang D. 2019. The complete plastid genome of *Zenia insignis* Chun (Leguminosae). *Mitochondrial DNA B Resour.* 4(2):2926–2927.
- Lee C, Choi IS, Cardoso D, de Lima HC, de Queiroz LP, Wojciechowski MF, Jansen RK, Ruhlman TA. 2021. The chicken or the egg? Plastome evolution and an independent loss of the inverted repeat in papilionoid legumes. *Plant J.* 107(3):861–875.
- Liu Y, Huang P, Li CH, Zang FQ, Zheng YQ. 2019. Characterization of the complete chloroplast genome of *Dalbergia cultrata* (Leguminosae). *Mitochondrial DNA B Resour.* 4(2):2369–2370.
- Marques A, Moraes L, Aparecida Dos Santos M, Costa I, Costa L, Nunes T, Melo N, Simon MF, Leitch AR, Almeida C, et al. 2018. Origin and parental genome characterization of the allotetraploid *Stylosanthes scabra* Vogel (Papilionoideae, Leguminosae), an important legume pasture crop. *Ann Bot.* 122(7):1143–1159.
- McMahon M, Hufford L. 2004. Phylogeny of Amorpheae (Fabaceae: Papilionoideae). *Am J Bot.* 91(8):1219–1230.
- Munz PA. 1959. A California flora. Berkeley (CA): University of California Press.
- Oyebanji O, Zhang R, Chen SY, Yi TS. 2020. New insights into the plastome evolution of the Millettoid/Phaseoloid clade (Papilionoideae, Leguminosae). *Front Plant Sci.* 11:151.
- Schattnner P, Brooks AN, Lowe TM. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res.* 33:686–689.
- 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.
- Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K, et al. 1986. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *Embo J.* 5(9):2043–2049.
- Straub SC, Doyle JJ. 2014. Molecular phylogenetics of *Amorpha* (Fabaceae): an evaluation of monophyly, species relationships, and polyploid origins. *Mol Phylogenet Evol.* 76:49–66.
- The Legume Phylogeny Working Group, Bruneau A, Doyle JJ, Herendeen P, Hughes C, Kenicer G, Lewis G, Mackinder B, Pennington RT, Sanderson MJ, Wojciechowski MF, et al. 2013. Legume phylogeny and classification in the 21st century: Progress, prospects and lessons for other species-rich clades. *Taxon.* 62:217–248.
- Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017. GeSeq - versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* 45(W1):W6–W11.
- Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ. 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* 44(W1):W232–W235.
- Wang J, Li Y, Li C, Yan C, Zhao X, Yuan C, Sun Q, Shi C, Shan S. 2019. Twelve complete chloroplast genomes of wild peanuts: great genetic resources and a better understanding of *Arachis* phylogeny. *BMC Plant Biol.* 19(1):504.
- Wicke S, Schneeweiss GM, dePamphilis CW, Müller KF, Quandt D. 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Mol Biol.* 76(3–5):273–297.
- Wilbur RL. 1975. A revision of the North American genus *Amorpha* (Leguminosae-Psoraleae). *Rhodora.* 77:337–409.
- Zhang J, Li Y, Heng S, Wang Y. 2020. The complete chloroplast genome sequence of *Dalbergia oliveri*. *Mitochondrial DNA B Resour.* 5(1):707–708.
- Zhang R, Wang YH, Jin JJ, Stull GW, Bruneau A, Cardoso D, De Queiroz LP, Moore MJ, Zhang SD, Chen SY, et al. 2020. Exploration of plastid phylogenomic conflict yields new insights into the deep relationships of Leguminosae. *Syst Biol.* 69(4):613–622.