





PLASTOME ANNOUNCEMENT



The complete plastome of *Nonea vesicaria* (L.) Rchb. (Boraginaceae), the first chloroplast genome belonging to the *Nonea* genus

Inês Carvalho Leonardo^{a,b} , Maria Teresa Barreto Crespo^{a,b} , Jorge Capelo^{c,d}  and
Frédéric Bustos Gaspar^{a,b} 

^aiBET, Instituto de Biologia Experimental e Tecnológica, Oeiras, Portugal; ^bITQB-NOVA, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Oeiras, Portugal; ^cECOCHANGE, CIBIO-InBIO – Research Centre in Biodiversity and Genetic Resources, Universidade do Porto, Vairão, Portugal; ^dINIAV, Instituto Nacional de Investigação Agrária e Veterinária I.P., Quinta do Marquês, Oeiras, Portugal

ABSTRACT

The predominantly Western Mediterranean weed *Nonea vesicaria* (L.) Rchb. can be found in agricultural or other man-made environments. Despite containing some beneficial compounds, extracts from this plant have also been described as detrimental and should be carefully monitored. In this study, the complete chloroplast of *N. vesicaria* isolate BPTPS250 is described, being the first available plastome from an isolate belonging to the *Nonea* genus. The chloroplast genome is 151,099 bp in length with a 37.3% GC content. It displays a quadripartite structure that contains a pair of inverted repeat regions (27,012 bp) that separate a large single-copy region (80,041 bp) and a small single-copy region (17,034 bp). A total of 134 genes were predicted, including 89 protein-coding genes, 37 tRNA genes, and 8 rRNA genes. The phylogenetic analysis confirmed the placement of *N. vesicaria* under the Boraginaceae family, belonging to the Boraginales order, with a close relationship with *Borago officinalis* L. This study will contribute to conservation, phylogenetic, and evolutionary studies, as well as DNA barcoding applications for food and feed safety and quality.

ARTICLE HISTORY

Received 21 February 2022
Accepted 23 June 2022

KEYWORDS

Nonea vesicaria;
Boraginaceae; complete
chloroplast genome;
Illumina MiSeq sequencing;
phylogenetic analysis

Nonea Medik. (monkswort) is a genus of herbaceous flowering plants (angiosperms) of the *Boraginoideae* subfamily, part of the *Boraginaceae* family, that encompasses ca. 42 genera with large flowers, with faucal and basal scales, gynobasic style, and often incurved roughly ovoidal nutlets (Chacón et al. 2016). *Nonea* includes ca. 45 species distributed in North Africa and from the Temperate Eurasia to the Indian Subcontinent. Although the *Nonea* centers of diversity are the Pontic-Caucasian and Irano-Turanic mountains, about 35 species occur in the Mediterranean basin, of which nine are found in its European part (WFO 2022).



Nonea vesicaria (L.) Rchb. (Reichenbach 1832) (red monkswort) is an annual or biennial greyish-green hairy weed found in Western Mediterranean's agricultural and other man-made environments. Its natural habitat is the Iberian Peninsula, Balearic Islands, and Sicily, and it is considered a sporadic alien in the U.S.A. and Canada (GBIF Secretariat 2022). *Nonea vesicaria* has been reported as beneficial in the phytoremediation of oil-polluted soils (Panchenko et al. 2022). Despite harboring stearidonic acid-rich seeds (an omega-3 fatty acid) and being a source of gamma-linolenic acid used for premenstrual syndrome, arthritis, blood pressure, and skin disorders, *N. vesicaria* extracts have been

described as having cytotoxic, hemolytic, and antioxidant activities (Mouffouk et al. 2020), and as such should be carefully monitored.

The material of *N. vesicaria* analyzed, isolate BPTPS250, was collected from a wild population in Oeiras municipality in Portugal (Collection date: 2021-05-20; Location: 38.70022 N 9.31913 W; Supplemental material Figure S1) with a specimen being conserved at the LISE Herbarium (INIAV, Oeiras, Portugal; Jorge Capelo: jorge.capelo@iniav.pt) under the voucher LISE: 96328 (Identified by: Jorge Capelo; Supplemental material Figure S2).

Total genomic DNA was extracted from young leaves, frozen in liquid nitrogen immediately after collection and kept at -80°C , using an adaptation of the Doyle and Doyle (1987) methodology. The obtained DNA was sheared by sonication using a Bioruptor (Diagenode), libraries were prepared with the NEBNext Ultra II DNA Library Prep Kit (New England Biolabs), and 150 bp paired-end sequencing was performed on an Illumina MiSeq platform using a v2 chemistry kit.

High-quality reads were used to assemble the complete chloroplast genome (sequence coverage: 112 \times) using the GetOrganelle pipeline (v1.7.4.1) (Jin et al. 2020), following the typical recipe suggested for Embryophyta plant plastome

CONTACT Frédéric Bustos Gaspar  fgaspar@ibet.pt  Food & Health Division, Instituto de Biologia Experimental e Tecnológica, Apartado 12, Oeiras 2781-901, Portugal

 Supplemental data for this article is available online at <https://doi.org/10.1080/23802359.2022.2095233>

© 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.

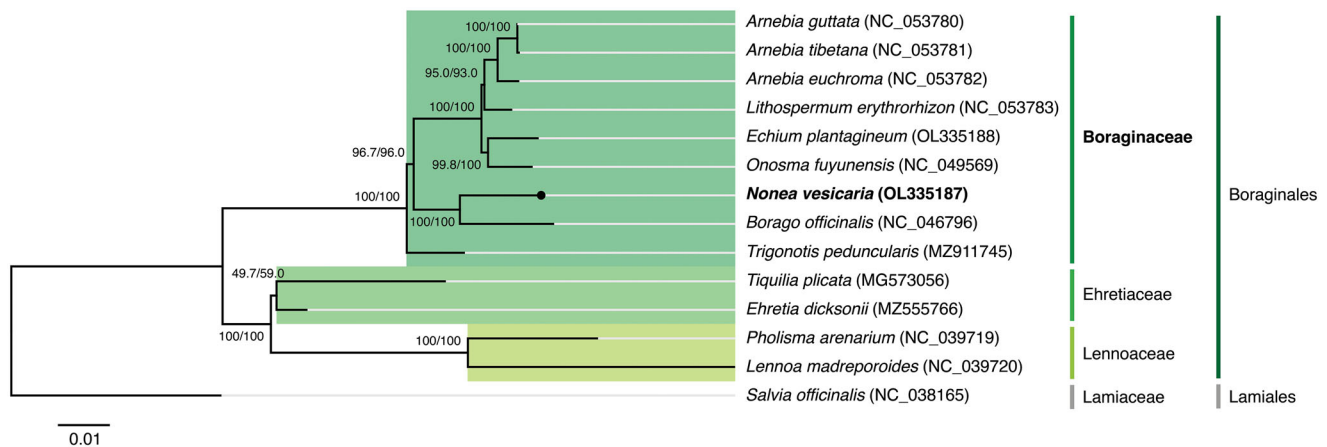


Figure 1. Maximum-likelihood tree inferred from the sequences coding for the shared proteome (25 coding sequences) of *Nonea vesicaria* isolate BPTPS250 and all 12 verified and complete chloroplast genomes belonging to the Boraginales order available in GenBank (Accession date: 2022.01.10). Numbers attached to the branches show the SH-aLRT and the UFBoot2 per cent supports (SH-aLRT/UFBoot2). *Salvia officinalis* (Lamiaceae) was used as the outgroup.

assembly (<https://github.com/Kinggerm/GetOrganelle>; see Supplemental material for additional details). The plastome annotation was performed using the GeSeq (Tillich et al. 2017) and CPGAVAS2 (Shi et al. 2019) tools with a subsequent manual curation using Geneious Prime 2022.0.1 (<https://www.geneious.com>; see Supplemental material for additional details).

The chloroplast genome of *N. vesicaria* isolate BPTPS250 (GenBank accession number: OL335187; Supplemental material Figure S3) is 151,099 bp in length with a 37.3% GC content, displaying a quadripartite structure that contains a pair of inverted repeat (IR) regions (27,012 bp, GC content 42.5%), separated by a large single-copy (LSC) region (80,041 bp, GC content 35.2%) and a small single-copy (SSC) region (17,034 bp, GC content 30.9%). A total of 134 genes were predicted, including 37 tRNA genes, 8 rRNA genes, and 89 protein-coding genes.

The phylogenetic analysis (see Supplemental material for additional details) was performed using the concatenated sequences coding for the shared proteome (25 coding sequences) extracted from all 12 verified and complete chloroplast genomes belonging to the Boraginales order available in GenBank (Accession date: 2022-01-10) and from the complete chloroplast genome of *N. vesicaria* obtained in this study. The sequences were aligned using MAFFT v7.450 (Katoh and Standley 2013) and further analyzed with the IQ-TREE 2 software package (Minh et al. 2020). The best-fit substitution model (TVM + F + R2 chosen according to the Bayesian Information Criterion) was selected according to ModelFinder (Kalyaanamoorthy et al. 2017), followed by a tree reconstruction (Figure 1) using IQ-TREE (Nguyen et al. 2015) using ultrafast bootstrap with UFBoot (10,000 replicates) (Hoang et al. 2018). The outgroup was *Salvia officinalis* L. (NC_038165) from the Lamiaceae family belonging to the Lamiales order.

The maximum likelihood tree showed that *N. vesicaria* is placed under the Boraginaceae family, belonging to the Boraginales order, and has an unsurprising close relationship with *Borago officinalis* L. (Chacón et al. 2016). Both are likely to belong to the same Boragineae tribe as sister clades, with 100/100 per cent support (SH-aLRT/UFBoot2). The phylogenetic analysis performed with the alignments of the Boraginales

complete chloroplast genomes also supports the same relationship between *N. vesicaria* and *B. officinalis*.

This complete chloroplast genome will contribute to conservation, phylogenetic, and evolutionary studies. It will also support DNA barcoding applications for food and feed safety and quality purposes that target the detection of species with cytotoxic and hemolytic potential.

Author contributions

The authors had the following contribution to the paper: FBG – conception and design; JC – collection and taxonomic identification of the studied specimen; ICL and FBG – analysis and interpretation of the data; FBG – drafting of the paper; ICL, JC, MTBC, and FBG – critical revision for intellectual content; FBG – final approval of the version to be published. All authors agree to be accountable for all aspects of the work herein presented.

Ethical approval statement

The species described and studied in this manuscript is not under legal protection status, either by national or European Union legislation, namely the 92/43/CEE Directive. A careful nondestructive collection protocol for voucher sampling was followed to guarantee the full future reproductive viability of the studied plant population.

Disclosure statement

No potential competing or conflict of interests was reported by the author(s).

Funding

Conducted under the project “barcodingPTplants – DNA barcoding aromatic, medicinal, and condiment plants from Portugal” [02/SAICT/2017 n° 31074] financed by Programa Lisboa 2020, Portugal 2020, and the European Union, through the European Regional Development Fund (ERDF), and with the support of INCDF funded by FCT and ERDF under the project 01/SAICT/2016 n° 022153. Funding from INTERFACE Programme, through the Innovation, Technology and Circular Economy Fund (FITEC), is gratefully acknowledged, as well as iNOVA4Health [UIDB/04462/2020 and UIDP/04462/2020] and the Associate Laboratory LS4FUTURE [LA/P/0087/2020], programs financially supported by FCT-MCTES.

ORCID

Inês Carvalho Leonardo  <http://orcid.org/0000-0001-8516-2487>
 Maria Teresa Barreto Crespo  <http://orcid.org/0000-0002-5605-1419>
 Jorge Capelo  <http://orcid.org/0000-0002-7634-6927>
 Frédéric Bustos Gaspar  <http://orcid.org/0000-0002-2948-6230>

Data availability statement

The data that supports this study is openly available in GenBank of NCBI at <https://www.ncbi.nlm.nih.gov> under the accession number [OL335187](#). The associated BioProject, BioSample, and SRA numbers are [PRJNA792567](#), [SAMN22746546](#), and [SRR16690176](#), respectively.

References

- Chacón J, Luebert F, Hilger HH, Ovchinnikova S, Selvi F, Cecchi L, Williams CM, Hasenstab-Lehman K, Sutorý K, Simpson MG, et al. 2016. The borage family (Boraginaceae s.str.): a revised infrafamilial classification based on new phylogenetic evidence, with emphasis on the placement of some enigmatic genera. *Taxon*. 65(3):523–546.
- Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull*. 19(1):11–15.
- GBIF Secretariat. 2022. *GBIF Backbone Taxonomy*. <https://doi.org/10.15468/39omei>
- Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: improving the ultrafast Bootstrap approximation. *Mol Biol Evol*. 35(2):518–522.
- Jin J-J, Yu W-B, Yang J-B, Song Y, dePamphilis CW, Yi T-S, Li D-Z. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol*. 21(1):241.
- Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermini LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*. 14(6):587–589.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment Software Version 7: improvements in performance and usability. *Mol Biol Evol*. 30(4):772–780.
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol*. 37(5):1530–1534.
- Mouffouk S, Mouffouk C, Bensouici C, Haba H. 2020. *In vitro* cytotoxic effect, hemolytic, and antioxidant activities of the algerian species *Nonea vesicaria* Rchb. *Curr Bioact Compd*. 16(8):1197–1204.
- Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. 32(1):268–274.
- Panchenko L, Muratova A, Biktasheva L, Galitskaya P, Golubev S, Dubrovskaya E, Selivanovskaya S, Turkovskaya O. 2022. Study of *Boraginaceae* plants for phytoremediation of oil-contaminated soil. *Int J Phytorem*. 24(2):215–219.
- Reichenbach HGL. 1832. *Flora Germanica excursoria ex affinitate regni vegetabilis naturali disposita, sive principia synopseos plantarum in Germania terrisque in Europa media adiacentibus sponte nascentium culturarumque frequentius: Insunt plantae: Acroblastae et Phylloblastae* (Vol. 1). Cnobloch.
- 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.
- 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.
- WFO. 2022. *World Flora Online*. <http://www.worldfloraonline.org/>.