

Replication Data for: Draft genome sequence data of Nothopassalora personata from Argentina (Original data) Draft Genome of Cercosporidium personatum (syn. Nothopassalora personata) (Original data) Replication Data for: Draft genome sequence data of Nothopassalora personata from Argentina (Original data)

Journal Pre-proof

Draft genome sequence data of Nothopassalora personata, peanut foliar pathogen from Argentina

Joaquin H. Monguillot , Renee S. Arias , Valerie A. Orner ,
Alicia N. Massa , Victor S. Sobolev , Nelson Bernardi Lima ,
Juan Paredes , Claudio Oddino , Marcelo Carmona ,
Cinthia Conforto

PII: S2352-3409(24)00129-X
DOI: <https://doi.org/10.1016/j.dib.2024.110158>
Reference: DIB 110158



To appear in: *Data in Brief*

Received date: 22 December 2023
Revised date: 30 January 2024
Accepted date: 31 January 2024

Please cite this article as: Joaquin H. Monguillot , Renee S. Arias , Valerie A. Orner , Alicia N. Massa , Victor S. Sobolev , Nelson Bernardi Lima , Juan Paredes , Claudio Oddino , Marcelo Carmona , Cinthia Conforto , Draft genome sequence data of Nothopassalora personata, peanut foliar pathogen from Argentina, *Data in Brief* (2024), doi: <https://doi.org/10.1016/j.dib.2024.110158>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2024 Published by Elsevier Inc.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

ARTICLE INFORMATION

Article title

Draft genome sequence data of *Nothopassalora personata*, peanut foliar pathogen from Argentina

Authors

Joaquin H. Monguillot^{a,b}, Renee S. Arias^{*c}, Valerie A. Orner^c, Alicia N. Massa^c, Victor S. Sobolev^c, Nelson Bernardi Lima^{a,b,d}, Juan Paredes^{a,b}, Claudio Oddino^e, Marcelo Carmona^f, Cinthia Conforto^{a,b}

Affiliations

^aInstituto de Patología Vegetal, Centro de Investigaciones Agropecuarias, Instituto Nacional de Tecnología Agropecuaria, IPAVE-CIAP-INTA, Av. 11 de Septiembre 4755, Córdoba, Argentina

^bUnidad de Fitopatología y Modelización Agrícola, Consejo Nacional de Investigaciones Científicas y Técnicas, UFYMA-CONICET, Av 11 de Septiembre 4755, Córdoba, Argentina

^cUnited States Department of Agriculture (USDA), Agricultural Research Service (ARS), National Peanut Research Laboratory (NPRL), 1011 Forrester dr. S.E., Dawson, GA, USA.

^dFacultad de Ciencias Agrarias, UNCA, Av. Belgrano 300, Catamarca, Argentina

^eFacultad de Agronomía y Veterinaria, IMICO, Ruta Nacional 36 km 601, Río Cuarto, Córdoba, Argentina

^fFacultad de Agronomía, Universidad de Buenos Aires, Cátedra de Fitopatología, Av. S. Martín 4453, Buenos Aires, Argentina

Corresponding author's email address and Twitter handle

*Corresponding author. E mail: renee.arias@usda.gov

Keywords

Late leaf spot; *Arachis hypogaea*; *Cercosporidium personatum*; MAT locus; groundnut; foliar disease

Abstract

Late leaf spot (LLS) caused by the Ascomycete *Nothopassalora personata* (N.p.) (Syn. *Cercosporidium personatum*) is the main foliar disease of peanuts in Argentina and in peanut producing areas of the world, causing up to 70% yield losses. The extremely slow growth of this fungus in culture, that takes around one month to form a 1 cm colony (0.45 mm/day), and the lack of adequate young tissues from where to extract nucleic acids, have hindered genetic studies of this pathogen. Here, we report the first genome sequence of a *N. personata* isolate from South America, as well as genetic variants on its conserved genes, and the complete sequence of its mating-type locus MAT1-2 idiomorph. The *N. personata* isolate IPAVE 0302 was obtained from peanut leaves in Córdoba, Argentina. The whole genome sequencing of IPAVE 0302 was performed as paired end 150 bp NovaSeq 6000 and *de novo* assembled. Clean reads were mapped to the reference genome for this species NRRL 64463 and the genetic variants on highly conserved genes and throughout the genome were analyzed. Sequencing data were submitted to NCBI GenBank Bioproject PRJNA948451, accession number SRR23957761. Additional Fasta files are available from Harvard Dataverse (<https://doi.org/10.7910/DVN/9AGPMG> and <https://doi.org/10.7910/DVN/YDO3V6>). The data reported here will be the basis for the analysis of genetic diversity of the LLS pathogen of peanut in Argentina, information that is critical to make decisions on management strategies.

SPECIFICATIONS TABLE

Subject	Genomics
Specific subject area	Fungal Plant Pathogen Genomics
Data format	Genome Sequencing, Raw data, De novo Assembly, Filtered Reads,

	Analyzed sequences
Type of data	Tables, Figures
Data collection	<p><i>Nothopassalora personata</i> was isolated from peanut (<i>Arachis hypogaea</i>) leaves of cv. Granoleico sampled in 2019, in Rio Cuarto, Córdoba, Argentina. Whole genome was sequenced on Illumina NovaSeq 6000 at the Genomics Center of the Univ. of California Davis, CA. Data were processed using CLC Genomics Workbench 23.0.4 (Qiagen, Aarhus, Denmark). Clean reads ≥ 140 nt were de novo assembled, and the assembly used to detect the MAT1-2 idiomorph locus of IPAVE 0302. Clean reads were also mapped to the genome reference NRRL 64463 and genome-wide distribution of Single Nucleotide Polymorphisms (SNP) quantified.</p>
Data source location	<p><i>N. personata</i> isolate was collected from Cordoba, Argentina, 32° 24' 30.5028" S, 63° 42' 18.9468" W. Sequencing data were placed in public repositories: NCBI GenBank and Harvard Dataverse.</p>
Data accessibility	<p>Bioproject PRJNA948451, accession number SRR23957761. Repository name: National Center for Biotechnology Information (NCBI), and Harvard Dataverse Data identification number: (https://doi.org/10.7910/DVN/9AGPMG and of https://doi.org/10.7910/DVN/YDO3V6)</p> <p>Direct URL to data: https://www.ncbi.nlm.nih.gov/sra/SRR23957761</p>

VALUE OF THE DATA

- This is the first report of a whole genome sequence of *Nothopassalora personata* from South America.
 - This genome shows many genetic variants when compared to the isolate NRRL 64463 from the United States. This information is critical to study the population genetics and evolution of the pathogen, and relevant to the development of disease management strategies.
 - This is the first report of the complete MAT1-2 idiomorph sequence of *Nothopassalora personata*, and its putative recombination sites in this species, as it was detected on isolates IPAVE 0302 and NRRL 64463.
 - The MAT locus affects sexual reproduction in fungi, and therefore their genetic diversity. The data reported here will be critical for genetic studies that focus on the biology of the pathogen.
 - The slow growth of *N. personata* in laboratory has hindered the study of its biology and genetics.
 - The data described here provide substantial information for comparative genomics and evolution of the species.
-

BACKGROUND

Argentina is the first-largest peanut producing and exporting country in South America with 336,817 ha cultivated area concentrated in the central-south region of Cordoba province [1;2]. The crop is affected by two important aerial fungal diseases, early leaf spot (ELS) and late leaf spot (LLS) (*Passalora arachidicola* (Hori). LLS, etiological agent *Nothopassalora personata* Berk. &M.A. Curtis (Syn. *Cercosporidium personatum*, teleomorph *Mycosphaerella berkeleyi* W.A.Jenkins) is considered more aggressive, difficult to control and present world-wide. Frequent fungicide applications to control this pathogen can lead to the emergence of resistance, often associated with mutations in fungicide target genes. Having genome sequencing information is of vital importance as it allows the study of regions and

mechanisms associated with fungicide resistance for subsequent more effective pathogen management.

Currently, there is no available information on the genome of *N. personata* for any peanut-producing country in the Southern Hemisphere. The purpose of this work was to obtain a draft genome of an isolate, IPAVE 0302, from the peanut producing area from Argentina to be used as a research model to understand the biology of this pathogen and its potential genetic diversity which could affect the strategies of disease control.

DATA DESCRIPTION

The draft genome sequence of isolate IPAVE 0302 (accession number SRR23957761) is available at the National Center for Biotechnology Information (NCBI) under the BioProject number PRJNA948451. Statistics are summarized in **Table 1**. Variants were detected in conserved genes normally used in phylogenetics, including ribosomal RNA cistron, RNA polymerase II largest subunit (RPB1) and RNA polymerase II second largest subunit (RPB2), using the same parameters as for genome-wide variant detection. Three variants were found on the 18S rRNA, three on the 28S rRNA, and one variant each in RPB1 and RPB2, **Table 2**.

Table 1. Genomic features and de novo assembly statistics for *Nothopassalora personata* IPAVE 0302

Attribute	Value
Genome size	34 306 332 bp
Number of contigs	1979
Largest contig (bp)	656 997
Average contig length (bp)	17 312
Coverage sequencing	1548 X
Number of scaffolds	1979
N25 (bp)	145 344
N50 (bp)	67 924
N75 (bp)	29 655
Overall GC content	51.91 %

Table 2. Genetic variants observed on conserved genes of *Nothopassalora personata* IPAVE 0302 and the reference genome NRRL 64463 using a threshold of 35 % of the reads. Nucleotides in parentheses “()” were only observed at 15 % threshold.

Locus	Ribosomal RNA cistron							RPB1	RPB2
	18S	ITS1	5.8S	ITS2	28S	28S	28S		
Position (nt)	23	55-56	1852	2108	8098	8315	8502	4869	1377
NRRL 64463	C/G	---	---	(C/T)	---	---	---	---	---
IPAVE 0302	---	CG/GT	C/T	C/T	C/T	T/C	A/C	C/T	A/G

Mapping of IPAVE 0302 genome sequencing reads to the 1061 contigs of the reference genome NRRL 64463, followed by calculation of number of single nucleotide polymorphisms (SNP) per kilobase were calculated, results presented in **Figure 1**. Examination of the MAT locus (Contig_298 in NRRL 64463), and further alignment to the reference genome showed potential recombination of this locus, **Figure 2**.

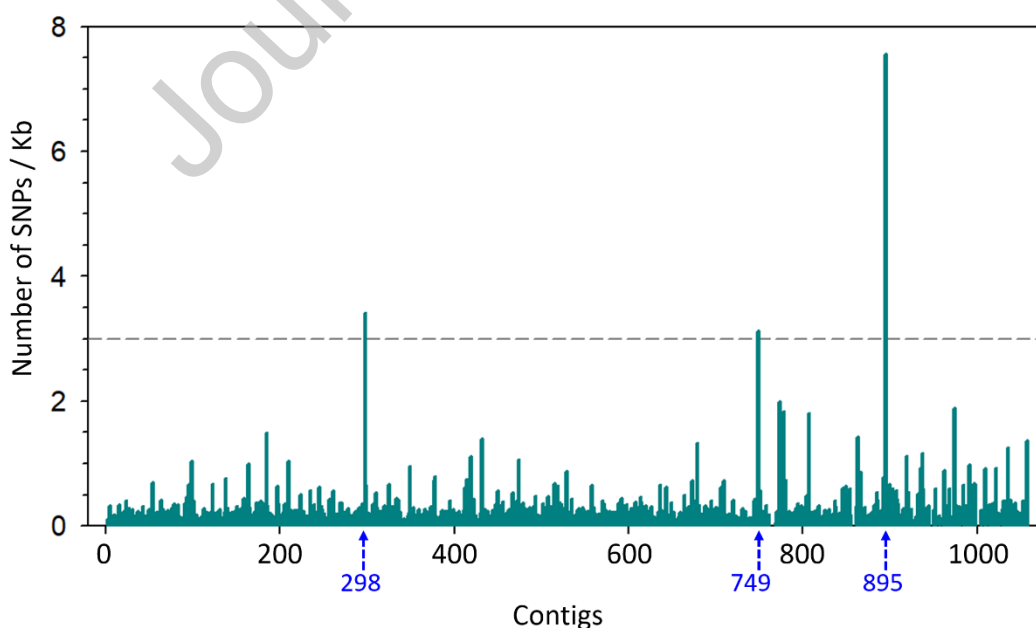


Figure 1. Genome-wide distribution of SNPs observed in IPAVE 0302 and not present in the reference genome, y-axis: SNPs per kilobase pair (kb), x-axis: contig numbers correspond to the location of SNPs using as reference the genome NRRL 64463.

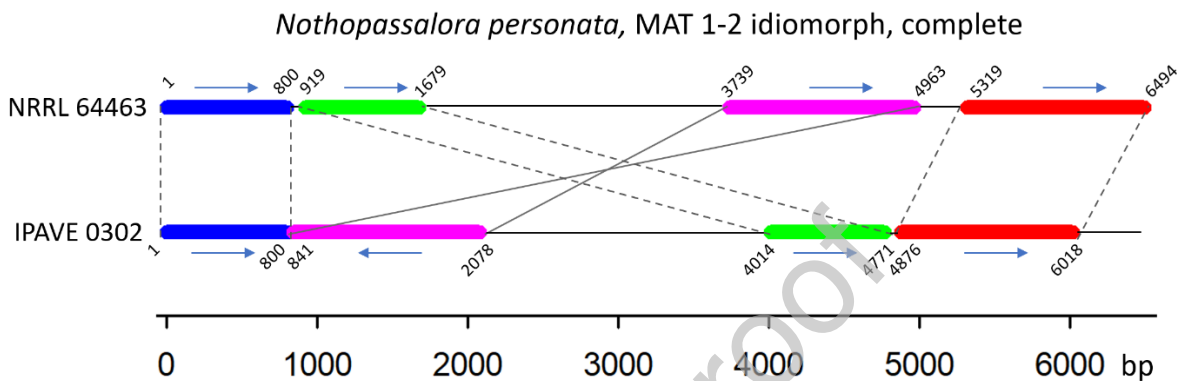


Figure 2. Schematic representation of the complete MAT1-2 idiomorph locus of *Nothopassalora personata* IPAVE 0302 compared to NRRL 64463, showing putative recombination.

EXPERIMENTAL DESIGN, MATERIALS AND METHODS

Fungal isolation

A monospore culture of *Nothopassalora personata* was isolated from symptomatic peanut field of Rio Cuarto, Cordoba, Argentina, in 2019. The leaves were surface disinfected by submersion in a 1.5% NaOCl solution for 45 s and washed twice with sterile distilled water. The stomata, conidiophores and conidia were transferred to a Petri dish containing water agar (WA) and spread on the culture medium using a Drigalsky spatula. Single conidia germinated when observed with a stereomicroscope were transferred to plates with POA (peanut leaf extract + oatmeal agar) + D (dextrose) [3]. The plates were incubated for five-ten days at 24 °C with a 12-h photoperiod. The isolate was preserved by Castellani's method

[4] and deposited in the Culture Collection of the Instituto de Patología Vegetal – IPAVE (Córdoba, Argentina).

Genomic DNA Preparation

Due to the slow growth of *N. personata*, genomic DNA was obtained from mycelium grown in liquid potato (200 g/L) + dextrose 10 g/L, (PD) medium + addition of 100 ml of the peanut leaf-extract (boiled 30 g peanut leaf cv. Granoleico/L) per 900 ml basal medium. After twelve days in a shaker incubator at 130 rpm, 24 °C with a 12-h photoperiod [2], DNA was extracted using the Wizard® Genomic DNA Purification Kit (Promega Corp.). The concentration of DNA was determined using NanoDrop spectrophotometer (Thermo Scientific, USA).

Genome sequencing, assembly, and annotation

Two genomic DNA libraries were prepared from the isolate IPAVE 0302 using Illumina TruSeq Prep Kit V2 (Illumina, San Diego, CA), and sequenced as paired end 150 base pairs (bp) in Illumina NovaSeq 6000 System (Illumina, San Diego, CA) at the University of California Davis Genome Center (Davis, CA). Clean reads were generated by removing potential adapters, low quality reads, and those shorter than 140 bp using CLC Genomics Work bench 23.0.4 (Qiagen, Aarhus, Denmark). Clean reads were mapped to the 8511 ribosomal RNA operon of *N.p.*, to the 5523 bp RNA polymerase II largest subunit gene (RPB1), and to the 4434 bp RNA polymerase II second largest subunit gene (RPB2) to corroborate taxonomic identification. A de novo whole-genome assembly was performed with clean reads \geq 140 bp of both libraries, in addition, clean reads were mapped to the 1061 contigs of the reference genome *Nothopassalora personata* (*N.p.*) isolate NRRL 64463. Genetic variants, including insertions, deletions, single and multi-nucleotide polymorphisms, and potential amino acid (AA) changes were detected and compared to NRRL 64463 genome wide as well as on rRNA cistron, RPB1, RPB2 [5]. All analyses were performed using CLC Genomics Workbench 23.0.4 (Qiagen, Aarhus, Denmark).

Mapping of IPAVE 0302 reads to the reference genome of *N.p.* NRRL 64463 showed the largest rates of genetic variants per kb in contig_298, contig_749 and contig_895, **Figure 1**. Contigs 749 and 895 are shorter than 3 kb, however contig_298 is 59.5 kb long and contains a ~6 kb region with homology to the mating-type locus MAT1-2 idiomorph of *Fulvia fulva*

strain Race5_Kim (Syn. *Cladosporium fulvum*, Syn. *Passalora fulva*) (GenBank Accession: CP090175) published by [6]. Further analysis of the putative MAT1-2 locus showed structural differences when compared to NRRL 64463 genome reference, with an apparent rearrangement of genes and possible recombination that resulted in an inversion, **Figure 2**. In other Ascomycetes, the MAT locus has been shown to play a role not only in sexual reproduction but also in pathogenesis [7]. In addition, sexual reproduction is a major source of genetic diversity of pathogens and can reduce the durability of host resistance [8]. Here, we report significant differences between the MAT1-2 locus of IPAVE 0302 isolated from Argentina and NRRL 64463 from U.S.A. The potential recombination in MAT1-2 locus will need to be considered when studying population genetics of *N. personata* in Argentina.

LIMITATIONS

Not applicable

ETHICS STATEMENT

The current work does not involve human subjects, animal experiments, or any data collected from social media platforms.

CRedit AUTHOR STATEMENT

Joaquin H Monguillot: Methodology, Writing –original draft. Renee Arias: Conceptualization, Writing –original draft, Formal analysis, Data curation and upload to repositories, Review and editing. Valerie Orner: Formal analysis, Data curation and upload to repositories. Massa Alicia: Formal analysis, Data curation and upload to repositories, Review and editing. Victor S. Sobolev: Writing –original draft. Nelson Bernardi Lima: Conceptualization, Methodology, Funding acquisition and resources, Review and editing. Juan A. Paredes: Methodology.

Claudio Oddino: Methodology, Funding acquisition and resources, Review and editing.
Marcelo Carmona: Funding acquisition and resources, Review and editing. Cinthia Conforto:
Conceptualization, Methodology, Supervision, upload to repositories, Writing –original draft,
Funding acquisition and resources, Project administration, Review and editing.

ACKNOWLEDGEMENTS

This work was supported by Fondo para la Investigacion Cientifica y Tecnologica (BID-PICT-2018-04057), grants from Instituto Nacional de Tecnologia Agropecuaria (2019-PE-E4-I086-001) and United States Department of Agriculture, Agricultural Research Service Research Project (6044-21000-005).

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

[1] D.F. Giordano, N.A. Pastor, L.F.M. Rouws, K.M. de Freitas, J.G. Erazo, A. del Canto, I. da Silva Coelho, C.M. Oddino, A.M. Torres

***Trichoderma harzianum* ITEM 3636 colonizes peanut roots as an endophyte and protects the plants against late leaf spot**

Symbiosis 89, 337–352 (2023). <https://doi.org/10.1007/s13199-023-00913-z>

[2] J.H. Monguillot, N.B Lima, J.A. Paredes, F.D, Giordano, C. Oddino, M.A. Rago, M. Carmona, C. Conforto

Morphological and molecular characterization of *Nothopassalora personata* from Argentina

J Plant Pathol. 105, 557–565 (2023). <https://doi.org/10.1007/s42161-022-01292-y>

[3] Y.A.M Abdou, W.E Cooper

Effect of Culture Medium and Light on Sporulation of Two Peanut Leaf Spotting Fungi, *Cercospora arachidicola* Hori and *Cercosporidium personatum*(Beck and Curtis) Deighton

Peanut Sci 1:11–14. (1974).

[4] A. Castellani.

Viability of some pathogenic fungi in distilled water

J Trop Med & Hyg. (1939), 42: 225–226.

[5] R.S. Arias, J.T. Dobbs, J.E. Stewart, E.G. Cantonwine, V.A. Orner, V.S. Sobolev, M.C. Lamb, A.N. Massa.

First draft genome and transcriptome of *Cercosporidium personatum*, causal agent of late leaf spot disease of peanut

BMC Res Notes. 16 (1):58. (2023) doi: 10.1186/s13104-023-06331-0. PMID: 37085927; PMCID: PMC10122357.

[6] A. Z. Zaccaron, Li-Hung Chen, A. Samaras and I. Stergiopoulos.

A chromosome-scale genome assembly of the tomato pathogen *Cladosporium fulvum* reveals a compartmentalized genome architecture and the presence of a dispensable chromosome.

Microbial Genomics 8:000819 (2022). DOI: 10.1099/mgen.0.000819[8]

[7] Q. Zheng, R. Hou, J. Zhang, J. Ma, Z. Wu, G. Wang, C. Wang, Jin-Rong Xu.

The MAT locus genes play different roles in sexual reproduction and pathogenesis in *Fusarium graminearum*.

Plos One 8(6): e66980 (2013). DOI: 10.1371/journal.pone.0066980

[8] B.A._McDonald, C. Linde C.

Pathogen population genetics, evolutionary potential, and durable resistance.

Annu Rev Phytopathol. 40:349-79 (2002) doi: 10.1146/annurev.phyto.40.120501.101443.

DECLARATION OF COMPETING INTERESTS

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Journal Pre-proof