Complete genome sequence of a novel alphanucleorhabdovirus infecting cassava (Manihot esculenta Crantz) in Brazil

Dayla Geovana Pereira Bezerra
UFRPE: Universidade Federal Rural de Pernambuco
Odaiza Fabiana Gomes Ferreira
UFRPE: Universidade Federal Rural de Pernambuco
José Ailton Cruz Macêdo dos Santos
UFRPE: Universidade Federal Rural de Pernambuco
Eder Jorge de Oliveira
EMBRAPA Centro Nacional de Pesquisa de Mandioca e Fruticultura Tropical
Paolo Margaria
DSMZ: Leibniz Institut - Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH
Rosana Blawid
rosana.blawid@ufrpe.br

Universidade Federal Rural de Pernambuco https://orcid.org/0000-0003-1001-7191

Research Article

Keywords:

Posted Date: September 16th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-7529671/v1

Version of Record: A version of this preprint was published at Archives of Virology on November 11th, 2025. See the published version at https://doi.org/10.1007/s00705-025-06463-7.

Abstract

The genus *Alphanucleorhabdovirus* comprises plant-infecting viruses with an unsegmented, negative-sense, single-stranded RNA genome. In this study, we report the complete genome sequence of a novel alphanucleorhabdovirus infecting cassava (*Manihot esculenta* Crantz), identified in the accession BGM-1275 from the germplasm collection at Embrapa Mandioca e Fruticultura (Cruz das Almas, Bahia, Brazil). The virus genome sequence is 13,625 nucleotides (nt) in length and displays the typical organization of alphanucleorhabdoviruses, with six predicted open reading frames. Pairwise comparisons of the whole genome sequence with other known alphanucleorhabdoviruses revealed nt identities in the range 31.0–67.8%, below the species demarcation threshold (75%). Phylogenetic analysis on the L gene confirmed placement of the novel virus within the genus *Alphanucleorhabdovirus*, with cassava alphanucleorhabdovirus 1 as closest relative. Reverse transcription-polymerase chain reaction assays detected the virus in symptomatic (cassava frogskin disease) and asymptomatic cassava samples collected in fields in Northeast Brazil (Paraíba and Rio Grande do Norte) and in other Embrapa accessions, suggesting widespread occurrence. The name “cassava alphanucleorhabdovirus 2” (CsRV2) is proposed for the newly discovered virus.

Full Text

The family *Rhabdoviridae* comprises a diverse group of viruses characterized by bacilliform virions with single-stranded, negative-sense RNA genomes, ranging from 10 to 16 kilobases in length. Members that infect plants and their arthropod vectors are classified within the subfamily *Betarhabdovirinae*, which currently includes twelve recognized genera: *Alphacytorhabdovirus*, *Alphagymnorhavirus*, *Alphanucleorhabdovirus*, *Betacytorhabdovirus*, *Betagymnorhavirus*, *Betanucleorhabdovirus*, *Deltanucleorhabdovirus*, *Dichorhavirus*, *Gammacytorhabdovirus*, *Gammanucleorhabdovirus*, *Trirhavirus*, and *Varicosavirus* [1]. Members of the genus *Alphanucleorhabdovirus* possess a genome with the five canonical open reading frames (ORFs) found in all members of the *Rhabdoviridae* family, encoding for the nucleocapsid protein (N), a phosphoprotein (P), a matrix protein (M), a glycoprotein (G), and a large protein (L, RNA-directed RNA polymerase), in the order 3'-N-P-M-G-L-5'. In addition, alphanucleorhabdoviruses encode a putative cell-to-cell movement protein (P3), typically located between the P and M genes, and eventually putative accessory proteins of unknown function, located either between N and P, or between G and L genes [2,3]. Transmission occurs in a persistent, circulative, and propagative manner via hemipteran insects [3], and can also occur through vegetative propagation or via infected sap [2]. In the context of a screening of cassava germplasm from Embrapa Mandioca e Fruticultura (Cruz das Almas, Bahia, Brazil), five accessions (BGM-1634, BGM-1618, BGM-0929, BGM-1275, BR-21GS-C3-130-17) showing virus symptoms were considered for a virome investigation study. Following total RNA extractions from leaf material of each accession, a ribosomal RNA-depleted library was prepared from a pool of RNA samples and sequenced on an Illumina NextSeq2000 platform to generate 59.069.502 million reads (2×150 bp). The subsequent bioinformatic analysis was conducted following the PhytoPipe workflow as described previously [4], with minor modifications, revealing, among others, a contig of 1,231 nt assigned to a putative novel alphanucleorhabdovirus, which is subject of this

study. Screening of the original, individual five plants by reverse transcription-polymerase chain reaction (RT-PCR) (Supplementary Table S1) allowed to detect the virus in accession BGM-1275, exhibiting typical symptoms of vein mosaic on leaves and of cassava frogskin disease (CFSD) [5] on tubers, which was thereafter considered for the prosecution of the study. Initially, primers (Supplementary Table S1) were designed based on assembled genomic sequence to amplify the nearly entire coding region of the genome with overlapping fragments, which were then subjected to Sanger sequencing. In parallel, newly-extracted RNA from BGM-1275 was used in another set of Illumina high-throughput sequencing (HTS) runs, including determination of genome ends using an RNA-tailing approach, which allowed assembly of the full-length genome sequence of the novel alphanucleorhabdovirus, consisting of 13,625 nt (acc. Number PX257435), and showing 99.8% identity with the genome sequence assembled from Sanger sequencing (not shown). The genome organization is consistent with members of the genus *Alphanucleorhabdovirus*, displaying six ORFs (Fig. 1), which were annotated in Geneious Prime (Dotmatics) v. 2025.1.2, with gene boundaries-confirmation through read coverage visualization, showing organization in the order 3'-N-P-P3-M-G-L-5'. The L gene spans nt positions 7,463 to 13,318 and corresponds to the most conserved region among rhabdoviruses, therefore it was considered for downstream analyses.

Analysis of the L ORF revealed the presence of conserved domains characteristic of viral RNA-dependent RNA polymerases (RdRp), including nucleotide binding-, template positioning-, and RNA synthesis-domain [6]. Specifically, six conserved motifs were identified: RxWGHP, Pre-motif A (GxxxKERE), Motif A (DFxKWNxxxR), Motif B (GxEGxRQKxWT), Motif C (GxGDNQ), and Motif D (GLPxKxxExWxSx$_7$Kx$_{13}$K), located at nucleotide positions 8,507–8,524, 9,035–9,058, 9,269–9,298, 9,488–9,520, 9,584–9,601, and 9,731–9,832, respectively. A BLASTx search of the L sequence revealed the highest amino acid identity (up to 75.9%) with the RdRp of cassava nucleorhabdovirus 1, followed by Ruellia alphanucleorhabdovirus 1, babaco nucleorhabdovirus 1, and potato yellow dwarf virus, with E-values ranging from 0.0 to 2.89e-90. The phylogenetic analysis of the L gene, including other plant-infecting rhabdoviruses, confirmed that the newly discovered virus belongs to the genus *Alphanucleorhabdovirus* (Fig. 2). Whole genome sequence (nt) alignment with sequences from other alphanucleorhabdoviruses available in the NCBI database (Supplementary Table S2) revealed identities ranging from 31.0% to 67.8% (Supplementary Table S3), lower than the threshold of 75% considered to assign viruses to different species [3]. Therefore, the determined sequence appears to belong to a novel alphanucleorhabdovirus, for which the name “cassava alphanucleorhabdovirus 2” (CsRV2) is proposed.

To conduct a small-scale survey, 22 additional cassava genotypes from Embrapa were analyzed by RT-PCR on RNA extracted from root material, including: BGM-0074, BGM-0179, BGM-0266, BGM-0544, BGM-0547, BGM-0560, BGM-0611, BGM-0671, BGM-1193, BGM-1294, BGM-1468, BGM-2102, BGM-2352, BR 11-34-64, BRS Caipira, BRS Poti Branca, BRS Tapioqueira, BGM-0065, BGM-1300, BGM-1618, BGM-0929, and BGM-1124. These genotypes showed characteristic symptoms of CFSD on the roots, such as size reduction, constrictions, grooves, and a corky epidermis. In addition, asymptomatic field samples from other municipalities in the Northeast Region of Brazil were also screened by RT-PCR for CsRV2: (i) a pool of 10 tuber samples of the *Rainha da Praia* variety, collected at Recife Supply and Logistics Center

(CEASA-Recife) and originating from fields in Rio Grande do Norte; (ii) a pool of 10 tubers of the same variety collected in Pitimbu, Paraíba; and (iii) tubers from genotypes BR-21GS-C3-130-17 and BGM-1634 from the Embrapa collection (Cruz das Almas, Bahia). All samples, both symptomatic (showing CFSD symptoms) and asymptomatic, tested positive for the novel alphanucleorhabdovirus.

Historically, nucleorhabdovirus particles were first reported in cassava leaves of asymptomatic plants using electron microscopy [7]. In this study, we assembled the complete genome sequence of a cassava alphanucleorhabdovirus, 45 years after the first observation of rhabdovirus-like particles in this plant. Our results suggest that the virus is widely distributed in cassava crops in Northeast Brazil. Although the impact of this virus on the crop is still poorly understood, CsRV2 infection does not appear to be directly associated with CFSD. Indeed recently, through sentinel experiments and HTS, Jimenez and coauthors reported that only torradoviruses, specifically cassava torrado-like virus (CsTLV), were consistently detected in roots expressing CFSD symptoms, supporting that single-infections by torradoviruses were sufficient to cause the disease [8].

Declarations

Funding This work was partially funded by the CAPES-PROBRAL Program, Call No. 09/2023 (Process No. 88881.895122/2023-01); the National Council for Scientific and Technological Development (CNPq, Bolsa PQ C); and the FACEPE APQ Call No. 29/2022 (Process No. APQ-1106-5.01/22). In addition, it was supported by the German Federal Ministry of Education and Research (BMBF) and conducted within the framework of the DAAD (German Academic Exchange Service), project number 57705342 to PM.

AuthorcontributionsD.G.P.B., O.F.G.F., J.A.C.M.S., E.J.O., and R.B. contributed to the study conception and design, material preparation, and data collection. Data analysis was conducted by D.G.P.B., O.F.G.F., P.M and R.B.. The first draft of the manuscript was written by D.G.P.B., and all authors commented and reviewed the manuscript. All authors read and approved the final manuscript.

Data availability The virus genome sequence was deposited in the GenBank database under accession number PX257435. The datasets generated during and in the current study are available from the corresponding author on reasonable request.

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Ethical approval Ethical approval does not apply.

The responsibility for the content of this publication lies with the authors only.

References

1. Kuhn JH, Abe J, Adkins S, Alkhovsky SV, Avšič-Županc T, Ayllón MA et al (2023) Annual(2023) taxonomic update of RNA-directed RNA polymerase-encoding negative-sense RNA viruses (realm

Riboviria: kingdom Orthornavirae: phylum Negarnaviricota). J Gen Virol 104:001864. https://doi.org/10.1099/jgv.0.001864

2. Jackson AO, Dietzgen RG, Goodin MM, Bragg JN, Deng M (2025) Biology of Plant Rhabdoviruses. Annu Rev Phytopathol 43:623–660. https://doi.org/10.1146/annurev.phyto.43.011205.141136
3. Walker PJ, Freitas-Astúa J, Bejerman N, Blasdell KR, Breyta R, Dietzgen RG et al (2022) ICTV Virus Taxonomy Profile: Rhabdoviridae 2022. J Gen Virol 103:001689. https://doi.org/10.1099/jgv.0.001689
4. Hu X, Hurtado-Gonzales OP, Adhikari BN, French-Monar RD, Malapi M, Foster JA et al (2023) PhytoPipe: a phytosanitary pipeline for plant pathogen detection and diagnosis using RNA-seq data. BMC Bioinformatics 24:470. https://doi.org/10.1186/s12859-023-05589-2
5. Pardo JM, Alvarez E, Becerra Lopez-Lavalle LA, Olaya C, Leiva AM, Cuellar WJ (2022) Cassava Frogskin Disease: Current Knowledge on a Re-Emerging Disease in the Americas. Plants 11:1841. https://doi.org/10.3390/PLANTS11141841/S1
6. Hu J, Miao T, Que K, Rahman MS, Zhang L, Dong X et al (2023) Identification, molecular characterization and phylogenetic analysis of a novel nucleorhabdovirus infecting Paris polyphylla var. yunnanensis. Sci Rep 13:10040. https://doi.org/10.1038/s41598-023-37022-2
7. Kitajima EW, Costa AS (1979) Rhabdovirus-like particles in tissues of five different plant species. Fitopatol Bras 4:55–62
8. Jimenez J, Caicedo S, Pardo JM, Gil-Ordóñez A, Alvarez-Quinto R, Mollov D et al (2024) Single torradovirus infections explain the mysterious cassava frogskin disease in the Americas. Sci Rep 14:29648. https://doi.org/10.1038/s41598-024-81142-2
9. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Syst Biol 59:307–321. https://doi.org/10.1093/sysbio/syq010
10. Letunic I, Bork P (2019) Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res 47:W256–W259. https://doi.org/10.1093/nar/gkz239

Figures

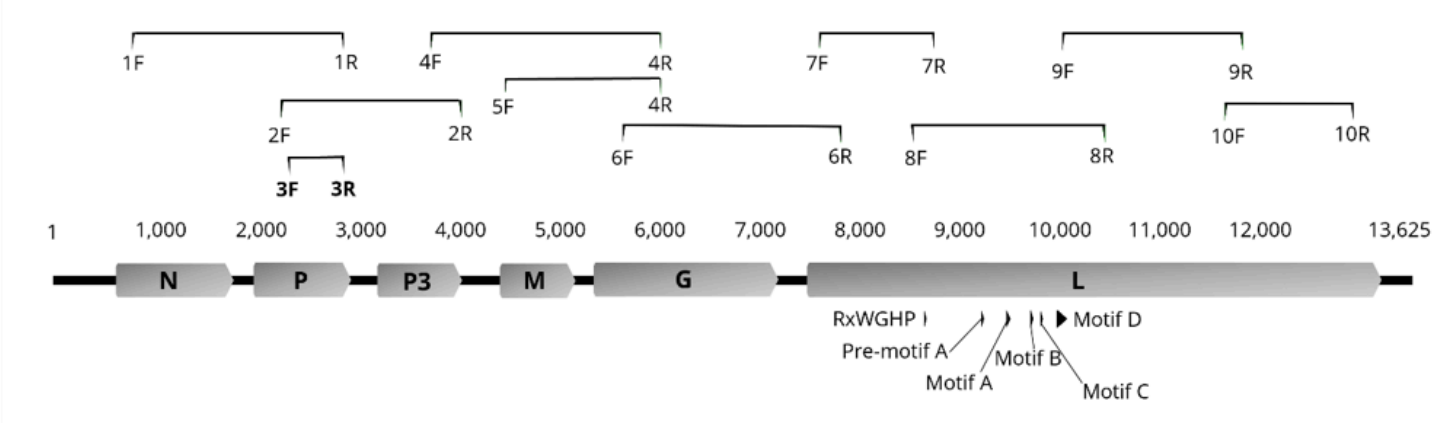

Figure 1

Genome organization of the novel cassava alphanucleorhabdovirus from this study. Boxes indicate the positions of the open reading frames (ORFs), and arrows represent conserved protein motifs within the L protein. ORFs: N, nucleocapsid protein; P, phosphoprotein; P3, movement protein; M, matrix protein; G, glycoprotein; L, replicase protein. Arrows indicate the primers used for genome validation by Sanger sequencing; the primer pair in bold was used in reverse transcription-polymerase chain reactions for virus detection.

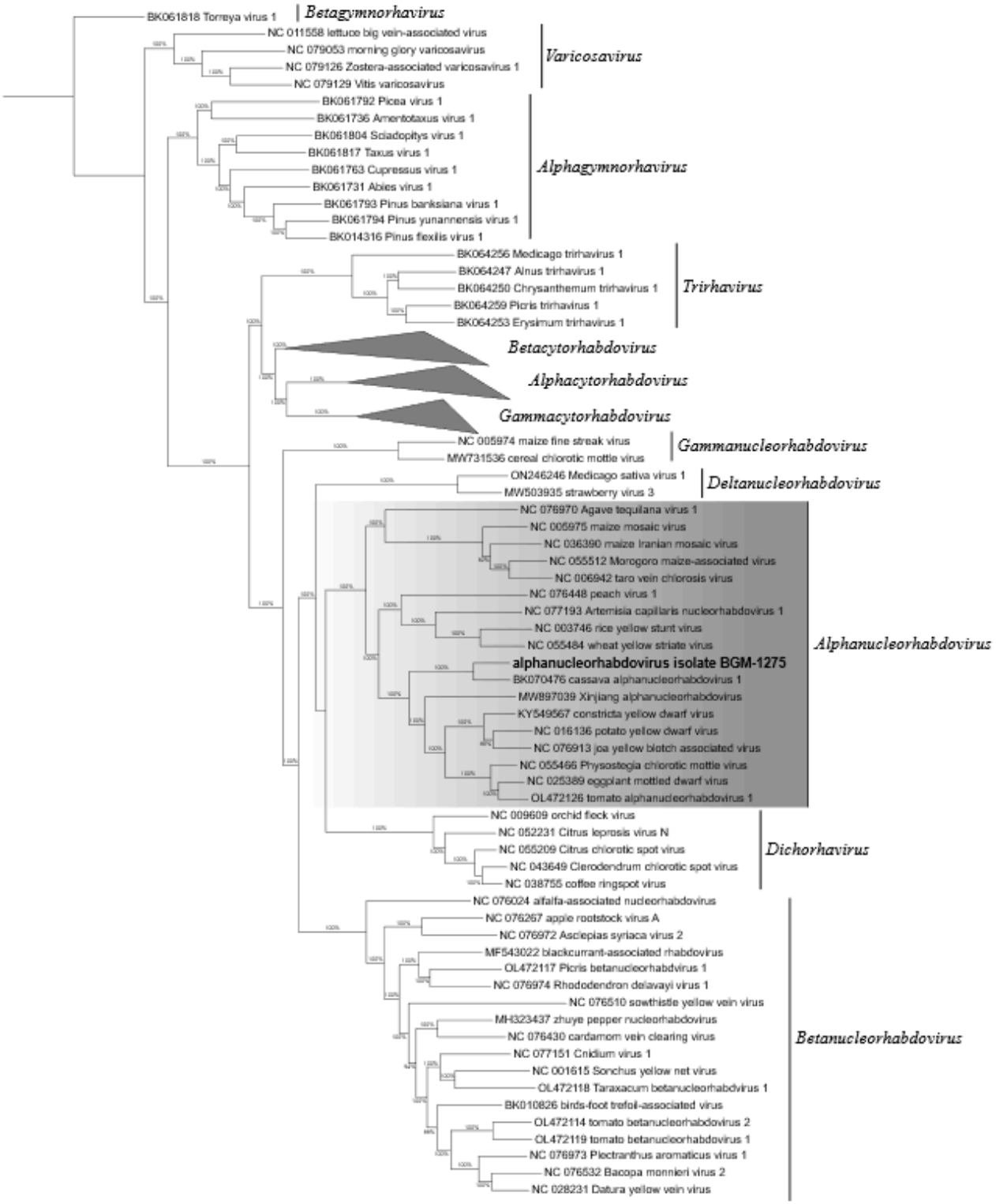

Figure 2

Unrooted phylogenetic tree constructed on the L gene according to the maximum likelihood method. The GTR+Γ (gamma-distributed rate variation) substitution model was identified as the best-fitting model. Branch support values (%) were estimated using the approximate Bayes method (aBayes). Colored rectangles represent different genera within the family. Virus sequences are labeled with their GenBank accession numbers and names. The novel alphanucleorhabdovirus (acc. Number PX257435) is highlighted in bold. Phylogenetic analyses were carried out using PhyML 3.0 [9], and the resulting trees were visualized and edited using the iTOL v.4 online platform [10].

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Submission2999360.txt
- SupplementaryTable1.docx
- SupplementaryTable2.docx
- SupplementaryTable3.docx