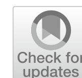
Tropical Plant Pathology (2022) 47:807–813
https://doi.org/10.1007/s40858-022-00536-7

SHORT COMMUNICATION



Badnavirus sequence diversity reveals one previously uncharacterized virus associated with air yam (*Dioscorea bulbifera* L.) in Brazil

Giancarlo B. L. Santos[1,2] · Mayra M. M. Ferro[2] · Frederico M. Feijo[2] · Roberto Ramos-Sobrinho[2] · Iraildes P. Assunção[2] · Gaus S. A. Lima[2]

Received: 11 April 2022 / Accepted: 7 October 2022 / Published online: 25 October 2022
© The Author(s), under exclusive license to Sociedade Brasileira de Fitopatologia 2022

Abstract

Dioscorea bulbifera L., commonly known as air yam, is an edible crop belonging to the botanical family *Dioscoreaceae*, and it has increasingly attracted attention due to its socioeconomical relevance. To molecularly characterize the species diversity of badnaviruses associated with *D. bulbifera* in Brazil, plant samples ($n=60$) were collected from different growing regions. Partial sequences of the reverse transcriptase (RT) and ribonuclease H (RNase H) domains were obtained from 26 PCR-positive samples. Based on pairwise nucleotide sequence comparisons, the RT-RNase H sequences (~580 bp in length) reported here belong to the badnaviral species dioscorea bacilliform AL virus (DBALV, $n=2$), dioscorea bacilliform SN virus (DBSNV, $n=8$), and dioscorea bacilliform TR virus (DBTRV, $n=2$), and the endogenous pararetrovirus dioscorea rotundata endogenous virus eDBV12 ($n=5$). Furthermore, a putative novel badnavirus tentatively named *Dioscorea bacilliform BL virus* ($n=9$) was partially characterized, sharing the highest nucleotide identity (75.7–79.9%) with eDBV12. The Bayesian phylogenetic tree reconstructed based on partial RT-RNase H nucleotide sequences showed the newly characterized isolates were clustered in three different clades, with dioscorea bacilliform BL virus (DBBLV) being more closely related to eDBV12 and forming a sister group with DBALV isolates. Together, these results reinforce the high badnaviral species diversity usually observed associated with *Dioscorea* spp. and constitute the first report of DBTRV in *D. bulbifera*.

Keywords *Dioscoreaceae* · *Caulimoviridae* · *Dioscorea* bacilliform viruses · EPRVs

Dioscorea bulbifera L. is an herbaceous twining vine belonging to the botanical family *Dioscoreaceae* and it is commonly called as air yam. The genetic diversity centers of *D. bulbifera* were mostly located in Asia and Africa, but it is currently spread throughout the Americas (Hammer 1998; Govaerts et al. 2007; Maurin et al. 2016). Air yam produces edible underground tubers and aerial bulbils that can be easily reproduced by vegetative propagation (Croxton et al. 2011). *Dioscorea bulbifera* is also classified into the category of non-conventional edible plants (Kinupp and Lorenzi 2014). In Brazil, crops belonging to the genus *Dioscorea* are socioeconomically important, being mainly cultivated by smallholder farmers, and are considered an alternative source for food security (Ferreira 2011).

Viruses belonging to the genus *Badnavirus* (family *Caulimoviridae*) have semicircular, double-stranded DNA (dsDNA) genomes of 7.0–9.2 kbp in size, encapsidated in a non-enveloped bacilliform particle, and encode three to seven open reading frames (ORFs) (Teycheney et al. 2020). The ORF1-encoded protein has been reported as virion-associated (Cheng et al. 1996), while ORF2 encodes a protein that binds to nucleic acid (Jacquot et al. 1996). The badnaviral ORF3 is the largest coding region, representing almost 80% of the viral genome, and encodes a polyprotein that includes important conserved domains such as the viral capsid, movement protein, aspartate protease, reverse transcriptase (RT), and ribonuclease H (RNase H) (Teycheney et al. 2020). Badnaviruses replicate through an RNA intermediate molecule, being referred to as plant pararetroviruses, and are mainly transmitted by mealybugs (*Pseudococcidae*) or, in some instances, aphids,

✉ Roberto Ramos-Sobrinho
robertorsobrinho@gmail.com

✉ Gaus S. A. Lima
gaus@ceca.ufal.br

[1] Campus Coruripe, Instituto Federal de Alagoas, Coruripe, Alagoas 57230-000, Brazil

[2] Campus de Engenharias e Ciências Agrárias, Universidade Federal de Alagoas, Rio Largo, Alagoas 57100-000, Brazil

in a non-circulative, semipersistent manner (Teycheney et al. 2020; Bhat et al. 2016).

Badnavirus-like particles have been first described in *Dioscorea* spp. in the 1970s (Harrison and Roberts 1973; Mantell and Haque 1979). Subsequently, partial and complete genome sequences obtained from *Dioscorea alata*, originally collected in Nigeria, have shown the presence of a new badnaviral species currently known as *Dioscorea bacilliform AL virus* (Phillips et al. 1999; Briddon et al. 1999). A second badnavirus species affecting *Dioscorea sansibarensis* has been later characterized, at the genome level, and it is presently referred to as *Dioscorea bacilliform SN virus* (Seal and Muller 2007). These badnaviruses are widespread in the main yam cultivation areas, indiscriminately infecting different *Dioscorea* species (Eni et al. 2008; Kenyon et al. 2008; Bousalem et al. 2009). *Dioscorea* bacilliform AL virus (DBALV) has been reported as the predominant badnaviral species infecting these hosts in northeastern Brazil (Guimarães et al. 2015). At least six other badnaviruses associated with *Dioscorea* spp. are officially accepted by the International Committee on Taxonomy of Viruses (ICTV), *Dioscorea bacilliform AL virus 2*, *Dioscorea bacilliform ES virus*, *Dioscorea bacilliform RT virus 1*, *Dioscorea bacilliform RT virus 2*, *Dioscorea bacilliform RT virus 3*, and *Dioscorea bacilliform TR virus* (https://talk.ictvonline.org/ictv-reports). Also, yam plants belonging to the *Dioscorea cayenensis-rotundata* species complex have been reported to harbor different groups of endogenous pararetroviral sequences (EPRVs) (Bousalem et al. 2009; Seal et al. 2014; Umber et al. 2014).

Here, the species diversity of badnaviruses associated with *D. bulbifera* from different growing regions in Brazil was assessed by PCR amplification and Sanger sequencing of the RT-RNase H domains. At least three previously reported badnaviruses were found, DBALV, *Dioscorea* bacilliform SN virus (DBSNV), and *Dioscorea* bacilliform TR virus (DBTRV). Sequences belonging to a putative novel badnavirus tentatively named dioscorea bacilliform BL virus (DBBLV) were recovered. Furthermore, badnavirus-like endogenous sequences were characterized, being most closely related to *Dioscorea rotundata* endogenous pararetroviruses.

Most bulbils of air yam were collected in northeastern Brazil, while some plants were obtained from the north, central-western, south, and southeastern regions of the country. To establish a gene bank collection, the air yam bulbils were planted at the experimental field of the Federal University of Alagoas, Rio Largo, Alagoas state, Brazil. Then, leaf samples were collected from each symptomatic and asymptomatic plant and kept at − 80 °C until being analyzed.

Total DNA was individually extracted from 100 to 200 mg of frozen leaf tissue using the method described by Doyle and Doyle (1987) and used as template for PCR amplification. The degenerated primers BadnaFP (5′-ATGCCITTYGGIITIAARAAYGCICC-3′) and BadnaRP (5′-CCAYTTRCAIACISCICCCCAICC-3′), which amplify a fragment of ~ 580 bp comprising part of the RT-RNase H domains of *Badnavirus* species (Yang et al. 2003), were used for virus detection. The PCR reactions were performed in a total volume of 15 µL, containing 1.5 µL of 10 × buffer (100 mM KCl, 100 mM Tris–HCl pH 9.0, 1% Triton-X), 1.2 µL of 2.5 mM dNTPs, 0.4 µL 50 mM $MgCl_2$, 0.2 µL of Taq DNA polymerase (Thermo Fisher Scientific, Carlsbad, CA, USA), 1.0 µL of each oligonucleotide (10 µM), 1.0 µL (50 ng) of total DNA, and 8.7 µL of nuclease-free water. The amplification conditions were initial denaturation at 94 °C for 4 min, 35 cycles of denaturation at 94 °C for 30 s, annealing at 50 °C for 30 s, and extension at 72 °C for 1 min, and a final extension step at 72 °C for 10 min. The PCR products were analyzed in 1% agarose gel, stained with ethidium bromide and visualized under ultraviolet light. Expected size amplicons (~ 580 bp) were gel-purified using the GFX™ PCR DNA and Gel Band Purification Kit (GE Healthcare, Illinois, USA) according to the manufacturer's protocol and directly Sanger sequenced with both BadnaFP and BadnaRP primers (Yang et al. 2003).

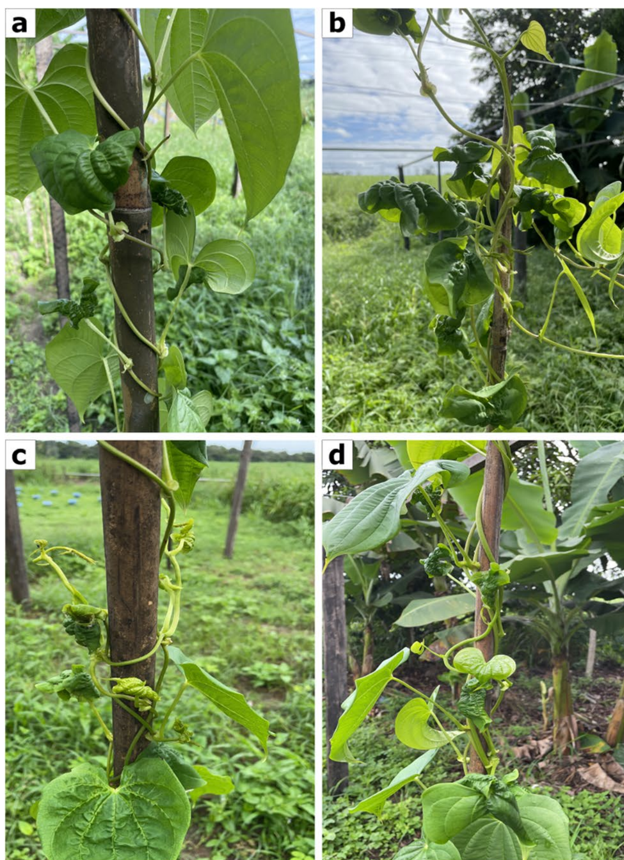
The contigs corresponding to the RT-RNase H nucleotide sequences were assembled and ambiguous sites were manually edited using CodonCode Aligner v. 4.1.1 (www.codoncode.com). The consensus sequences were initially analyzed with the BLASTn algorithm (Altschul 1990) to identify their closest matches among virus sequences available in the NCBI non-redundant GenBank database (https://www.ncbi.nlm.nih.gov/genbank). Then, similar sequences obtained from GenBank (Supplementary Table S1) were used for species demarcation of the new isolates via pairwise nucleotide sequence comparisons using Sequence Demarcation Tool v. 1.2 (Muhire et al. 2014). The *Badnavirus* species demarcation criterion of ≥ 80% nucleotide identity for the RT-RNase H domains established by the ICTV was adopted (Teycheney et al. 2020).

The RT-RNase H nucleotide sequences of the isolates reported here and badnaviral sequences retrieved from GenBank were aligned using the MUSCLE algorithm (Edgar 2004), and manually adjusted in MEGA 7.0 (Kumar et al. 2015). The phylogenetic relationship was determined by Bayesian inference (BI) through the CIPRES web portal (Miller et al. 2010) using MrBayes v.3.2.3 (Ronquist et al. 2012), assuming a general time reversible (GTR) nucleotide substitution model with a gamma (G) model of rate heterogeneity and invariable (I) sites, determined using MrModeltest 2.3 (Posada and Buckley 2004) according to the Akaike information criterion (AIC). The analysis consisted of two replicates with four chains each for 10 million generations and sampling every 1000 generations. The first 2500 trees per run were discarded as a burn-in.

The posterior probability values (Rannala and Yang 1996) were determined from the majority rule consensus tree reconstructed with the 15,000 remaining trees. The BI tree was edited in FigTree v.1.4 (ztree.bio.ed.ac.uk/software/figtree) and Inkscape (https://inkscape.org/pt/).

A total of 60 air yam bulbils were obtained from different growing regions in Brazil ($n_{north} = 4$; $n_{northeastern} = 47$; $n_{south} = 4$; $n_{southeastern} = 3$; $n_{central\text{-}western} = 2$). Leaf samples were collected from symptomatic and asymptomatic plants generated from the air yam bulbils (Fig. 1a–d) and individually tested, by PCR, using the BadnaFP/BadnaRP primer pair (Yang et al. 2003). Expected size amplicons of approximately 580 bp were observed from 39 of 60 plants, suggesting an incidence level of 65% (Supplementary Table S2). High incidence of badnaviruses, ranging from 72.0–93.3%, in *Dioscorea* spp. have been previously reported in Brazil (Lima et al. 2013; Guimarães et al. 2015; Nascimento et al. 2020), demonstrating that badnaviruses are widespread in commercial plantations and germplasm collections of yams. Although the PCR primers described by Yang et al. (2003) are well known for being unable to distinguish between episomal and integrated RT-RNase H sequences, it is an important and frequently used tool for badnavirus detection (Teycheney et al. 2020; Bhat et al. 2016; Luo et al. 2022). The amplification products from 26 PCR-positive samples were bidirectionally Sanger sequenced and showed that at least 12 of 26 plants were infected by badnaviral species officially accepted by the ICTV.

The isolates DBJ2 and DBJ3 shared 99.0% nucleotide identity with each other and showed highest identity (84.3–95.2%) with DBALV sequences, while the isolates DBMT1 and DBMT2 showed 99.8% identity between them and were more closely related to DBTRV, at 86.0–86.2% identity. The isolates DBAM1, DBJG, DBT6, DBC2, DBM6, DBM8, DBM9, and DBMU showed 97.9–99.4% nucleotide identity among them, and shared greater identity with DBSNV, at 82.3–83.2%. The isolates DB11, DBM123, DB31, DB32, DBB2, DBCO4, DBE, DBCU2, and DBH showed 82.6–99.1% nucleotide identity with one another, and shared highest nucleotide identity with



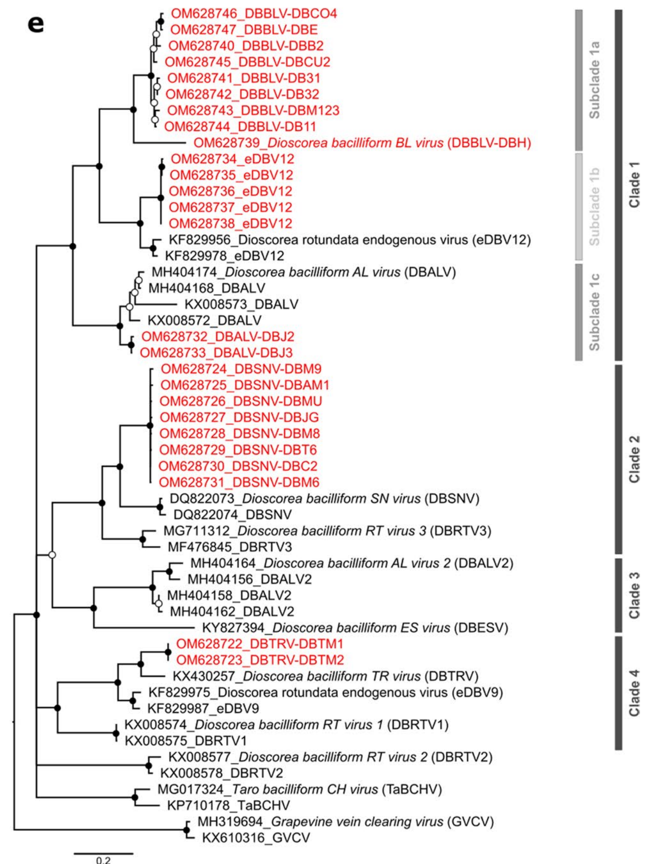


Fig. 1 *Dioscorea bulbifera* plants exhibiting foliar chlorosis and leaf curling symptoms. Partial RT-RNase H sequences of *Dioscorea bacilliform SN virus* (**a**, isolate DBMU), *Dioscorea bacilliform TR virus* (**b**, isolate DBMT2), *Dioscorea bacilliform BL virus* (**c**, isolate DBM123), and *Dioscorea rotundata* endogenous virus (**d**, isolate DBB1) were recovered, and their phylogenetic relationship was reconstructed using Bayesian inference (**e**). Posterior probability values between 0.95 and 1.0 (filled circles) and 0.50 and 0.94 (empty circles) are shown near to each branch node. The isolates reported here are indicated in red

Table 1 Percentages of pairwise nucleotide identity based on partial RT-RNase H sequences of the new isolates and badnaviral isolates retrieved from NCBI-GenBank

	DBALV2	DBESV	DBRTV2	DBSNV[1]	DBSNV[2]	DBRTV3	DBRTV1	DBTRV[1]	DBTRV[2]	eDBV9	TaBCHV	eDBV12[1]	eDBV12[2]	DBBLV	DBALV[1]	DBALV[2]	TaBV
DBALV2	89.0 – 100.0																
DBESV	70.8 – 71.6	100.0															
DBRTV2	64.4 – 66.5	64.6 – 67.2	94.0 – 100.0														
DBSNV[1]	69.3 – 71.2	69.3 – 69.7	68.4 – 71.2	97.4													
DBSNV[2]	69.9 – 74.2	68.4 – 69.1	67.8 – 69.5	82.3 – 83.2	97.9 – 99.4												
DBRTV3	69.7 – 71.7	66.3 – 68.6	69.0 – 71.7	76.4 – 79.4	78.7 – 80.4	90.6 – 100.0											
DBRTV1	67.0 – 70.2	68.5 – 68.9	71.0 – 72.7	71.0 – 71.3	71.4 – 74.4	69.5 – 70.2	99.1 – 99.4										
DBTRV[1]	69.7 – 71.3	68.4	65.4 – 65.5	67.0 – 67.4	67.4 – 68.4	65.9 – 66.3	73.2 – 73.6	100.0									
DBTRV[2]	67.8 – 68.6	66.1 – 66.3	64.5 – 67.6	66.3 – 67.8	66.1 – 67.6	66.7 – 67.1	71.3 – 71.5	86.0 – 86.2	99.8								
eDBV9	64.2 – 70.4	68.6 – 69.2	66.9 – 68.2	68.6 – 69.4	68.8 – 70.8	67.8 – 69.0	72.9 – 73.5	82.8 – 83.6	83.7 – 84.7	95.5							
TaBCHV	65.0 – 68.9	65.5 – 67.0	65.7 – 68.9	64.2 – 66.7	65.5 – 69.1	64.6 – 66.7	67.0 – 69.1	63.9 – 65.2	63.9 – 65.6	65.7 – 67.6	88.4 – 99.3						
eDBV12[1]	65.3 – 69.2	66.9 – 67.8	67.6 – 69.2	65.1 – 65.7	64.3 – 65.5	65.1 – 65.9	71.9 – 72.3	68.8 – 69.4	66.0 – 66.7	68.2 – 68.8	69.8 – 70.2	95.7					
eDBV12[2]	65.4 – 69.6	67.9 – 68.3	65.0 – 67.4	65.6 – 66.4	65.8 – 67.2	64.9 – 66.6	70.3 – 71.1	66.0 – 66.6	62.2 – 63.0	68.3 – 69.1	68.0 – 68.9	88.6 – 89.2	99.4 – 100.0				
DBBLV	67.2 – 70.7	67.6 – 68.9	66.5 – 70.1	67.0 – 70.9	64.6 – 70.2	66.5 – 70.5	70.0 – 71.3	66.5 – 69.1	66.1 – 69.4	69.0 – 74.5	67.0 – 71.2	75.7 – 79.9	73.4 – 77.5	82.6 – 99.1			
DBALV[1]	65.7 – 71.3	69.3 – 71.0	65.4 – 68.2	68.2 – 71.0	67.2 – 71.8	67.4 – 70.0	71.0 – 73.4	69.9 – 71.2	66.7 – 70.0	70.8 – 72.9	68.4 – 70.0	73.6 – 76.1	72.7 – 75.2	73.2 – 77.7	85.2 – 100.0		
DBALV[2]	69.1 – 70.5	69.5 – 69.9	66.9 – 68.7	69.1 – 70.1	69.5 – 71.1	68.5 – 69.7	71.7 – 72.1	69.9	66.7 – 67.4	71.7 – 72.5	69.1 – 70.1	73.9 – 75.1	73.6 – 74.6	74.1 – 76.5	84.3 – 95.2	99.0	
TaBV	61.3 – 66.0	61.8 – 64.1	59.3 – 61.4	60.5 – 62.0	62.7 – 68.4	62.0 – 65.3	61.7 – 64.8	61.7 – 62.0	61.7 – 62.9	60.9 – 64.8	62.4 – 66.9	61.4 – 62.5	60.2 – 63.1	59.9 – 65.0	61.6 – 67.7	63.0 – 65.2	86.1 – 99.8

DBALV, *Dioscorea* bacilliform AL virus; *DBALV2*, *Dioscorea* bacilliform AL virus 2; *DBESV*, *Dioscorea* bacilliform ES virus; *DBRTV1*, *Dioscorea* bacilliform RT virus 1; *DBRTV2*, *Dioscorea* bacilliform RT virus 2; *DBRTV3*, *Dioscorea* bacilliform RT virus 3; *DBSNV*, *Dioscorea* bacilliform SN virus; *DBTRV*, *Dioscorea* bacilliform TR virus; *DBBLV*, *Dioscorea* bacilliform BL vírus; *TaBV*, taro bacilliform virus; *TaBCHV*, taro bacilliform CH virus; *eDBV9*, *Dioscorea* rotundata endogenous virus; *eDBV12*, *Dioscorea* rotundata endogenous virus

[1]Sequences retrieved from GenBank

[2]Sequences reported in this study

Dioscorea rotundata endogenous pararetrovirus eDBV12, at 75.7–79.9% identity, suggesting these sequences may represent a putative new badnaviral species for which the name *Dioscorea bacilliform BL virus* is tentatively proposed (Table 1). Additional studies are needed to determine if this new species represents an episomal or integrated badnavirus. Finally, the isolates DB21, DBB1, DBG, DBT2, and DBT3 shared 99.4–100.0% with one another, and were identified as eDBV12, at 88.6–89.2% nucleotide identity (Table 1). The new sequences reported here were deposited in NCBI-GenBank under accession nos. OM628722-OM628747 (Table 2).

Badnaviruses are able to infect tropical and subtropical crops, including *Dioscorea* spp., of great socioeconomical importance worldwide and can lead to economic losses between 10 and 90% (Phillips et al. 1999; Briddon et al. 1999; Seal and Muller 2007; James et al. 2011; Eni et al. 2008; Kenyon et al. 2008; Bousalem et al. 2009; Silva et al. 2015; Deeshma and Bhat 2015; Bhat et al. 2016; Luo et al. 2022). Yam plants affected by badnaviruses usually exhibit disease symptoms such as leaf chlorosis and deformation, and dwarfism, which can lead to a reduction in the photosynthetic capacity of the infected plant with deleterious effects on production, tuber quality, and plant death (Thouvenel and Dumont 1988; 1990). Recently, a taxonomic positioning study in *Badnavirus* suggested partial RT-RNase H sequences (~580 bp) are sufficient for species demarcation (Ferreira et al. 2019). Therefore, based on the ICTV-approved ≥ 80% nucleotide identity species demarcation criterion for RT-RNase sequences into the genus *Badnavirus* (Teycheney et al. 2020), isolates of DBALV, DBSNV, DBTRV, and DBBLV were found to be largely spread in *D. bulbifera* growing areas in Brazil, reinforcing this host harbors a high badnaviral species diversity that can negatively impact the disease management. To our knowledge, this is the first report of DBTRV in *D. bulbifera* worldwide.

Endogenous pararetroviral sequences (EPRVs) have been shown to be integrated into the genome of the African yam, *D. cayenensis-rotundata* complex, but no evidence of EPRVs has been found in other yam species such as *D. alata* and *D. sansibarensis* (Bousalem et al. 2009; Seal et al. 2014; Umber et al. 2014). In the present study, EPRV sequences (eDBV12) previously reported in *D. cayenensis-rotundata* were also characterized from *D. bulbifera* plants using PCR primers amplifying the badnaviral RT-RNase H domains. These results emphasize the importance of EPRVs present

Table 2 Badnaviruses associated with *Dioscorea bulbifera* in Brazil

Isolate	Species	Acronym	Host	GenBank Accession #
DBMT1	*Dioscorea bacilliform TR virus*	DBTRV	*Dioscorea bulbifera*	OM628722
DBMT2	*Dioscorea bacilliform TR virus*	DBTRV	*Dioscorea bulbifera*	OM628723
DBM9	*Dioscorea bacilliform SN virus*	DBSNV	*Dioscorea bulbifera*	OM628724
DBAM1	*Dioscorea bacilliform SN virus*	DBSNV	*Dioscorea bulbifera*	OM628725
DBMU	*Dioscorea bacilliform SN virus*	DBSNV	*Dioscorea bulbifera*	OM628726
DBJG	*Dioscorea bacilliform SN virus*	DBSNV	*Dioscorea bulbifera*	OM628727
DBM8	*Dioscorea bacilliform SN virus*	DBSNV	*Dioscorea bulbifera*	OM628728
DBT6	*Dioscorea bacilliform SN virus*	DBSNV	*Dioscorea bulbifera*	OM628729
DBC2	*Dioscorea bacilliform SN virus*	DBSNV	*Dioscorea bulbifera*	OM628730
DBM6	*Dioscorea bacilliform SN virus*	DBSNV	*Dioscorea bulbifera*	OM628731
DBJ2	*Dioscorea bacilliform AL virus*	DBALV	*Dioscorea bulbifera*	OM628732
DBJ3	*Dioscorea bacilliform AL virus*	DBALV	*Dioscorea bulbifera*	OM628733
DBT2	Dioscorea rotundata endogenous virus	eDBV12	*Dioscorea bulbifera*	OM628734
DBB1	Dioscorea rotundata endogenous virus	eDBV12	*Dioscorea bulbifera*	OM628735
DB21	Dioscorea rotundata endogenous virus	eDBV12	*Dioscorea bulbifera*	OM628736
DBT3	Dioscorea rotundata endogenous virus	eDBV12	*Dioscorea bulbifera*	OM628737
DBG	Dioscorea rotundata endogenous virus	eDBV12	*Dioscorea bulbifera*	OM628738
DBH	*Dioscorea bacilliform BL virus*	DBBLV	*Dioscorea bulbifera*	OM628739
DBB2	*Dioscorea bacilliform BL virus*	DBBLV	*Dioscorea bulbifera*	OM628740
DB31	*Dioscorea bacilliform BL virus*	DBBLV	*Dioscorea bulbifera*	OM628741
DB32	*Dioscorea bacilliform BL virus*	DBBLV	*Dioscorea bulbifera*	OM628742
DBM123	*Dioscorea bacilliform BL virus*	DBBLV	*Dioscorea bulbifera*	OM628743
DB11	*Dioscorea bacilliform BL virus*	DBBLV	*Dioscorea bulbifera*	OM628744
DBCU2	*Dioscorea bacilliform BL virus*	DBBLV	*Dioscorea bulbifera*	OM628745
DBCO4	*Dioscorea bacilliform BL virus*	DBBLV	*Dioscorea bulbifera*	OM628746
DBE	*Dioscorea bacilliform BL virus*	DBBLV	*Dioscorea bulbifera*	OM628747

in the genome of yams for implementation of reliable molecular detection tools and, although no evidence of infectious EPRVs has been found in *Dioscorea* spp., it may represent a challenge for yam germplasm conservation and exchange of genetic materials between breeding programs.

The BI phylogenetic tree reconstructed based on partial RT-RNase H sequences showed that the new sequences were clustered in three different clades (Fig. 1e). The isolates reported here and sharing the highest nucleotide identity with *Dioscorea* endogenous sequences, were clustered into two sister subgroups, referred to as subclades 1a and 1b. The subclade 1a was comprised by isolates representing the putative new species DBBLV (isolates DB11, DBM123, DB31, DB32, DBB2, DBCO4, DBE, DBCU2, and DBH), while the isolates DB21, DBB1, DBG, DBT2, and DBT3 clustered in a monophyletic group with eDBV12 sequences (subclade 1b; Fig. 1e). These results agree with the pairwise sequence comparisons and reinforce that the sequences in subclade 1a may represent a new badnaviral species. However, since only partial sequences were obtained in the present study, and these isolates were more closely related to eDBV12 endogenous sequences, additional studies are necessary to clarify their episomal or integrated origin.

The isolates DBJ2 and DBJ3 grouped together with other DBALV sequences in the phylogenetic subclade 1c, while the DBSNV (isolates DBAM1, DBJG, DBT6, DBC2, DBM6, DBM8, DBM9, and DBMU) and DBTRV (isolates DBMT1 and DBMT2) sequences were placed in divergent phylogenetic groups, clades 2 and 4, respectively (Fig. 1e). So far, eight distinct *Dioscorea*-infecting badnaviruses have been characterized at the genome level (Briddon et al. 1999; Seal and Muller 2007; Bömer et al. 2016; Umber et al. 2016; Sukal et al. 2017; Bömer et al. 2018; Sukal et al. 2020), with DBALV and DBSNV being previously reported associated with *D. bulbifera* (Sukal et al. 2020; Nascimento et al. 2020). Although high badnaviral species diversity in *D. bulbifera* was observed here, additional samples from different growing regions must be analyzed, as well as the complete genomes need to be characterized. Also, assessing the extant species diversity of badnaviruses infecting *D. bulbifera*, and other *Dioscorea* species, and estimating the evolutionary mechanisms acting on diversification of these viruses are important steps to improve disease identification and management.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40858-022-00536-7.

Acknowledgements M.M.M.F. and F.M.F. were the recipients of CAPES postdoctoral fellowships. I.P.A. and G.S.A.L. wish to thank CNPq for the research fellowship.

Author contribution Conceptualization: Giancarlo B.L. Santos, Roberto Ramos-Sobrinho, and Gaus S.A. Lima; methodology: Giancarlo B.L. Santos, Mayra M.M. Ferro, and Frederico M. Feijo; formal analysis and investigation: Giancarlo B.L. Santos, Mayra M.M. Ferro, Frederico M. Feijo, and Roberto Ramos-Sobrinho; writing—original draft preparation: Giancarlo B.L. Santos, Mayra M.M. Ferro, and Frederico M. Feijo; writing—review and editing: Giancarlo B.L. Santos, Roberto Ramos-Sobrinho, Iraildes P. Assunção, and Gaus S.A. Lima; Funding acquisition: Iraildes P. Assunção and Gaus S.A. Lima; resources: Iraildes P. Assunção and Gaus S.A. Lima; supervision: Roberto Ramos-Sobrinho and Gaus S.A. Lima.

Data availability The nucleotide sequence datasets generated during the current study are available in the NCBI-GenBank repository, under accession numbers OM628722-OM628747.

Declarations

Conflict of interest The authors declare no competing interests.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology 215:403–410

Bhat A, Hohn T, Selvarajan R (2016) Badnaviruses: the current global scenario. Viruses 8:177

Bömer M, Turaki A, Silva G, Kumar P, Seal S (2016) A sequence-independent strategy for amplification and characterization of episomal badnavirus sequences reveals three previously uncharacterized yam badnaviruses. Viruses 8:188

Bömer M, Rathnayake AI, Visendi P, Silva G, Seal SE (2018) Complete genome sequence of a new member of the genus Badnavirus, Dioscorea bacilliform RT virus 3, reveals the first evidence of recombination in yam badnaviruses. Archives of Virology 163:533–538

Bousalem M, Durand O, Scarcelli N, Lebas BSM, Kenyon L, Marchand JL, Lefort F, Seal SE (2009) Dilemmas caused by endogenous pararetroviruses regarding the taxonomy and diagnosis of yam (*Dioscorea* spp.) badnaviruses: analyses to support safe germplasm movement. Archives of Virology 154:297–314

Briddon RW, Phillips S, Brunt A, Hull R (1999) Analysis of the sequence of *Dioscorea alata* bacilliform virus; comparison to other members of the badnavirus group. Virus Genes 18:277–283

Cheng CP, Lockhart BE, Olszewski NE (1996) The ORF I and II proteins of *Commelina yellow mottle virus* are virion-associated. Virology 223:263–271

Croxton MD, Andreu MA, Williams DA, Overholt WA, Smith JA (2011) Geographic origins and genetic diversity of air-potato (*Dioscorea bulbifera*) in Florida. Invasive Plant Science and Management 4:22–30

Deeshma KP, Bhat AI (2015) Complete genome sequencing of piper yellow mottle virus infecting black pepper, betelvine, and Indian long pepper. Virus Genes 50:172–175

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small amounts of fresh leaf tissue. Phytochemistry Bulletin 19:11–15

Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:1–19

Eni AO, Hughes JDA, Asiedu R, Rey MEC (2008) Sequence diversity among badnavirus isolates infecting yam (*Dioscorea* spp.) in Ghana, Togo, Benin and Nigeria. Archives of Virology 153:2263–2272

Ferreira AB (2011) Sistemas de cultivo do cará Dioscorea spp. por sistemas de cultivo do cará Dioscorea spp. Botucatu

Ferreira CHLH, Jordão LJ, Ramos-Sobrinho R, Ferro MMM, Silva SJC, Assunção IP, Lima GSA (2019) Diversification into the genus *Badnavirus*: phylogeny and population genetic variability. Ciência Agrícola 17:59–72

Govaerts R, Wilkin P, Saunders RMK (2007) World checklist of Dioscoreales. Yams and their allies. The Board of Trustees of the Royal Botanic Gardens, Kew, London, 1–65.

Guimarães KM, Silva SJC, Melo AM, Ramos-Sobrinho R, Lima JS, Zerbini FM, Assunção IP, Lima GSA (2015) Genetic variability of badnaviruses infecting yam (*Dioscorea* spp.) in northeastern Brazil. Tropical Plant Pathology 40:111–118

Hammer RL (1998) Diagnosis: *Dioscorea*. Wildland Weeds 2:8–10

Harrison B, Roberts I (1973) Association of virus-like particles with internal brown spot of yam (*Dioscorea alata*). Tropical Agriculture 50:335–340

Jacquot E, Hagen LS, Jacquemond M, Yot P (1996) The open reading frame 2 product of cacao swollen shoot badnavirus is a nucleic acid-binding protein. Virol 225:191–195

James AP, Geijskes RJ, Dale JL, Harding RM (2011) Development of a novel rolling-circle amplification technique to detect banana streak virus that also discriminates between integrated and episomal virus sequences. Plant Disease 95:57–62

Kenyon L, Lebas BSM, Seal SE (2008) Yams (*Dioscorea* spp.) from the South Pacific Islands contain many novel badnaviruses: implications for international movement of yam germplasm. Archives of Virology 153:877–889

Kinupp VF, Lorenzi H (2014) Plantas alimentícias não convencionais (PANC) no Brasil: guia de identificação, aspectos nutricionais e receitas ilustradas. Plantarum, Nova Odessa, 768.

Kumar S, Stecher G, Tamura K (2015) MEGA7: molecular evolutionary genetics analysis version 7. Molecular Biology Evolution 33:1870–1874

Lima JS, Lima ATM, Castillo-Urquiza GP, Silva SJC, Assunção IP, Michereff SJ, Zerbini FM, Lima GSA (2013) Variabilidade genética de isolados de badnavírus infectando inhame (*Dioscorea* spp.) no Nordeste do Brasil. Tropical Plant Pathology 38:349–353

Luo GF, Podolyan A, Kidanemariam DB, Pilotti C, Houliston G, Sukal AC (2022) A review of viruses infecting yam (*Dioscorea* spp.). Viruses 14:662

Mantell SH, Haque SQ (1979) Internal brown spot disease of yams. Yam Virus Project Bulletin (Trinidad Tobago) ill:1–13

Maurin O, Muasya AM, Catalan P, Shongwe EZ, Viruel J, Wilkin P, van der Bank M (2016) Diversification into novel habitats in the Africa clade of *Dioscorea* (Dioscoreaceae): erect habit and elephant's foot tubers. BMC Evolutionary Biology 16:238

Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Gateway Computing Environments Workshop (GCE). LA, Ieee, New Orleans, pp 1–8

Muhire BM, Varsani A, Martin DP (2014) SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. PLoS ONE 9:e108277

Nascimento JP, Oliveira ML, Peixinho GS, Ferro MMM, Silva SJC, Versey EA, Lima GSA, Assunção IP (2020) Incidência e caracterização molecular de Badnavírus em bancos de germoplasma de inhame no Brasil. Summa Phytopathologica 46:242–249

Phillips S, Briddon RW, Brunt AA, Hull R (1999) The partial characterization of a badnavirus infecting the greater asiatic or water yam (*Dioscorea alata*). Journal of Phytopathology 147:265–269

Posada D, Buckley T (2004) Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Systematic Biology 53:793–808

Rannala B, Yang ZH (1996) Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. Journal of Molecular Evolution 43:304–311

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. Systematics Biology 61:539–542

Seal S, Muller E (2007) Molecular analysis of a full-length sequence of a new yam badnavirus from *Dioscorea sansibarensis*. Archives of Virology 152:819–825

Seal S, Turaki A, Muller E, Kumar PL, Kenyon L, Filloux D, Galzi S, Lopez-montes A, Iskra-caruana M (2014) The prevalence of badnaviruses in West African yams (*Dioscorea cayenensis*-rotundata) and evidence of endogenous pararetrovirus sequences in thier genomes. Virus Research 186:144–154

Silva JM, Jobim LJ, Ramos-Sobrinho R, Lima JS, Assunção IP, Cruz MM, Lima GSA (2015) Incidence and species diversity of badnaviruses infecting sugarcane from a germplasm collection in Brazil. Tropical Plant Pathology 40:212–217

Sukal A, Kidanemariam D, Dale J, James A, Harding R (2017) Characterization of badnaviruses infecting *Dioscorea* spp. in the Pacific reveals two putative novel species and the first report of *Dioscorea* bacilliform RT virus 2. Virus Research 238:29–34

Sukal AC, Kidanemariam DB, Dale JL, Harding RM, James AP (2020) Characterization and genetic diversity of *Dioscorea* bacilliform viruses present in a Pacific yam germplasm collection. Plant Pathology 69:576–584

Teycheney PY, Geering AD, Dasgupta I, Hull R, Kreuze JF Lockhart B, Muller E, Olszewsk N, Pappu H, Pooggin MM, Richert-Pöggeler KR, Schoelz JE, Seal S, Stavolone L, Umber M (2020) ICTV Virus taxonomy profile: *Caulimoviridae*. Journal of General Virology 101:1025

Thouvenel JC, Dumont R (1989) An epidemiological approach to the study of yam mosaic viruses in the Ivory Coast. In: Symposium de l'ISTRC. 8. Ho Chi Minh-Ville: CIRAD-ISA-MAE, 5 p

Thouvenel JC, Dumont R (1990) Yield decreases in yam infected with mosaic virus in Cote d'Ivoire. L'agronomie Tropicale 45:125–129

Umber M, Filloux D, Muller E, Laboureau N, Galzi S, Roumagnac P, Pavis C, Teycheney P, Seal SE (2014) The genome of African yam (*Dioscorea cayenensis*-rotundata complex) hosts endogenous sequences from four distinct badnavirus species. Molecular Plant Pathology 15:790–801

Umber M, Gomez R, Gelabale S, Bonheur L, Pavis C, Teycheney P (2016) The genome sequence of *Dioscorea* bacilliform TR virus, a member of the genus *Badnavirus* infecting *Dioscorea* spp., sheds light on the possible function of endogenous *Dioscorea* bacilliform viruses. Archives of Virology 162:517–521

Yang IC, Hafner GJ, Dale JL, Harding RM (2003) Genomic characterization of taro bacilliform virus. Archives of Virology 148:937–949

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.