UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO

PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO

PROGRAMA DE PÓS-GRADUAÇÃO EM FITOPATOLOGIA

Tese de Doutorado

Diversidade viral em mandioca (*Manihot esculenta* Crantz.) nas regiões Norte e Nordeste do Brasil

Géssyka Rodrigues de Albuquerque

Recife – PE

2022

GÉSSYKA RODRIGUES DE ALBUQUERQUE

Diversidade viral em mandioca (*Manihot esculenta* Crantz.) nas regiões Norte e Nordeste do Brasil

Tese apresentada ao Programa de Pós-Graduação em Fitopatologia da Universidade Federal Rural de Pernambuco, como parte dos requisitos para obtenção do título de Doutora em Fitopatologia.

COMITÊ DE ORIENTAÇÃO:

Orientadora: Profa. Dra. Rosana Blawid

RECIFE-PE

JULHO – 2022

Dados Internacionais de Catalogação na Publicação
Universidade Federal Rural de Pernambuco
Sistema Integrado de Bibliotecas
Gerada automaticamente, mediante os dados fornecidos pelo(a) autor(a)

A345d Albuquerque, Gésssyka Rodrigues de
Diversidade viral em mandioca (Manihot esculenta Crantz.) nas regiões Norte e Nordeste do Brasil / Gésssyka Rodrigues de Albuquerque. - 2022.
120 f. : il.

Orientadora: Rosana Blawid.
Inclui referências, apêndice(s) e anexo(s).

Tese (Doutorado) - Universidade Federal Rural de Pernambuco, Programa de Pós-Graduação em Fitopatologia, Recife, 2022.

1. High-throughput sequencing (HTS). 2. diversidade viral em mandioca. 3. cassava torrado-like virus. 4. cassava virus X. 5. cassava polero-like virus. I. Blawid, Rosana, orient. II. Título

CDD 632

Diversidade viral em mandioca (*Manihot esculenta* Crantz.) nas regiões Norte e Nordeste do Brasil

GÉSSYKA RODRIGUES DE ALBUQUERQUE

Tese defendida e aprovada pela Banca Examinadora em: 27/07/2022.

ORIENTADORA:

_____

Profa. Dra. Rosana Blawid (UFRPE)

EXAMINADORES:

_____

Dra. Alessandra de Jesus Boari (Embrapa Amazônia Oriental)

_____

Profa. Dra. Elineide Barbosa de Souza (UFRPE)

_____

Dr. Fernando Lucas de Melo (OnSite Genomics)

_____

Prof. Dr. Leonardo Silva Boiteux (UFRPE)

RECIFE-PE

JULHO-2022

À minha mãe, **Profa. Delazy Albuquerque*, pelo amor e incentivo***
DEDICO

“Até aqui nos ajudou o Senhor”

I Samuel 7:12

AGRADECIMENTOS

A Deus pela presença constante em minha vida.

À minha mãe, minha grande incentivadora, Delazy Albuquerque e ao meu pai, Ebenézer Albuquerque e ao meu irmão Odon Neto, pelo amor incondicional. Pelas inúmeras viagens Aliança-Recife para me acompanhar. Ao meu Noivo Lucas Lucena por todo apoio e incentivo dado ao longo desses anos. Vocês foram essenciais.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de Doutorado.

À Universidade Federal Rural de Pernambuco (UFRPE) e ao Programa de Pós-Graduação em Fitopatologia, pela formação no curso de Doutorado em Fitopatologia.

À minha orientadora, Profa. Dra. Rosana Blawid pela confiança, pelas oportunidades que me foram dadas, assim como pelas contribuições e auxílio indispensável para o desenvolvimento deste trabalho e minha formação profissional.

À Pesquisadora Dra., Alessandra de Jesus Boari da Embrapa Amazônia Oriental, ao Pesquisador Dr. Eder Jorge de Oliveira da Embrapa Mandioca e Fruticultura, aos Pesquisadores, Dr. Stephan Winter e Dr. Paolo Margaria do Instituto DSMZ na Alemanha, por compartilhar conhecimentos importantes para o desenvolvimento deste trabalho e todo apoio dado.

Aos queridos do Laboratório de Fitovirologia, pelos momentos compartilhados e todo o apoio ao longo desses anos. Em especial, José Ailton, Carlos Henrique, Elayni Araújo e Marcelo Gonçalves.

SUMÁRIO

RESUMO GERAL

A mandioca (*Manihot esculenta* Crantz.) é uma planta da família Euphorbiaceae nativa do Brasil. A cultura é uma importante fonte de subsistência nas regiões tropicais, e o Brasil é atualmente o sexto maior produtor mundial. No entanto, as doenças da mandioca representam um desafio e podem impactar amplamente a sua produção, e entre os principais agentes fitopatogênicos estão os vírus. As principais viroses encontradas na mandioca no Brasil são o mosaico-das-nervuras, causada pelo CsVMV (cassava vein mosaic virus), e o mosaico-comum causado pelo CsCMV (cassava common mosaic virus), que podem ser responsáveis por perdas de até 30% na produção brasileira. Uma vez que a metagenômica vem contribuído para o desenvolvimento da virologia de plantas, esta tese teve como objetivo estudar a diversidade viral em amostras de (meta)transcriptomas de mandioca das regiões Norte e Nordeste do Brasil. Um total de 28 bibliotecas de amostras de mandioca do banco de germoplasma da Embrapa Mandioca e Fruticultura foram analisadas por sequenciamento de alto desempenho (*High-throughput Sequencing*, HTS) pela plataforma disponível no Departamento de Vírus de Plantas do Instituto Leibniz DSMZ (Coleção Alemã de Microrganismos e Cultura de Células, Alemanha). Também foi sequenciado um pool de amostras de mandioca do banco de germoplasma da Embrapa Amazônia Oriental. Duas espécies de torradovírus foram encontradas, uma agrupando-se com o isolado de CsTLV da Colômbia, e uma nova espécie formando um clado distinto, a qual chamamos de cassava torrado-like virus 2 (CsTLV-2). Este vírus também foi encontrado nas amostras de mandioca do banco de germoplasma da Embrapa Amazônia Oriental mostrando a dispersão e importância dos torradovírus em plantações no Brasil. Nesta tese também foi relatada pela primeira vez no Brasil a presença do cassava virus X (CsVX), associado a um satélite de potexvírus. Essa sequência brasileira compartilha identidade de 82,9 (nt) e 94,9% (aa) na região da ORF1 e 82,3 (nt) e 97,2% (aa) nas regiões da CP com a sequência venezuelana de CsVX (número acc: NC_034375). Adicionalmente foi montado o genoma completo do cassava polero-like virus a partir dos dados de HTS e confirmado por sequenciamento Sanger. Este polerovírus foi detectado em três pools de amostras de tecido foliar de mandioca do banco de germoplasma da Embrapa Amazônia Oriental. Desta forma, o presente trabalho contribui para ampliar o conhecido sobre a diversidade viral em amostras de mandioca das regiões Norte e Nordeste do Brasil.

Palavras-chaves: *High-throughput sequencing* (HTS); diversidade viral em mandioca; cassava torrado-like virus; cassava virus X; cassava polero-like virus.

GENERAL ABSTRACT

Cassava (*Manihot esculenta* Crantz.) belongs to the Euphorbiaceae family and it is a plant species native to Brazil. The crop is an important source of subsistence in tropical regions, and Brazil is currently the sixth largest producer in the world. However, cassava diseases pose a challenge and can greatly impact their production, and among the main phytopathogenic agents are viruses. The main viruses found in cassava in Brazil are cassava vein mosaic virus (CsVMV), and cassava common mosaic virus (CsCMV), which can be responsible for losses of up to 30% in the Brazilian production. Since metagenomics has contributed to the development of the area of plant virology, this thesis aimed to study viral diversity in samples of cassava (meta)transcriptomics from the North and Northeast regions of Brazil. A total of 28 libraries of cassava samples from the germplasm bank of Embrapa Cassava and Fruits were analyzed by high-throughput Sequencing (HTS) performed by a platform available at the Plant Virus Department of the Leibniz Institute DSMZ (German Collection of Microorganisms and Cell Culture, Germany). A pool of cassava samples from the germplasm bank of Embrapa Eastern Amazon was also sequenced. A pool of cassava samples from the germplasm bank of Embrapa Eastern Amazon was also sequenced by HTS. Two species of torradovirus were found, one grouping with the CsTLV isolate from Colombia, and a new species forming a distinct clade, which we named cassava-like virus 2 (CsTLV-2). This virus was also found in samples from the germplasm bank of Embrapa Eastern Amazon, showing the dispersion and importance of torradovirus in cassava plantations in Brazil. In the present thesis, it was also reported for the first time in Brazil the presence of cassava virus X (CsVX), associated with a potexvirus satellite. This Brazilian sequence shares 82.9 (nt) and 94.9% (aa) sequence identity in the ORF1 region and 82.3 (nt) and 97.2% (aa) in the CP regions with the Venezuelan CsVX sequence (acc. number: NC_034375). Additionally, the complete genome of the cassava polero-like virus was assembled from the HTS data and confirmed by Sanger sequencing. This polerovirus was detected in three pools of cassava leaf tissue samples from the germplasm bank of Embrapa Eastern Amazon. Therefore, the present work contributes to expand the knowledge about the viral diversity in cassava samples from the North and Northeast regions of Brazil.

Keywords: *High-throughput sequencing* (HTS); viral diversity in cassava plants; cassava torrado-like virus; cassava virus X; cassava polero-like virus

CAPÍTULO I

Introdução Geral

DIVERSIDADE VIRAL EM MANDIOCA (*Manihot esculenta* Crantz.) NAS REGIÕES NORTE E NORDESTE DO BRASIL

INTRODUÇÃO GERAL

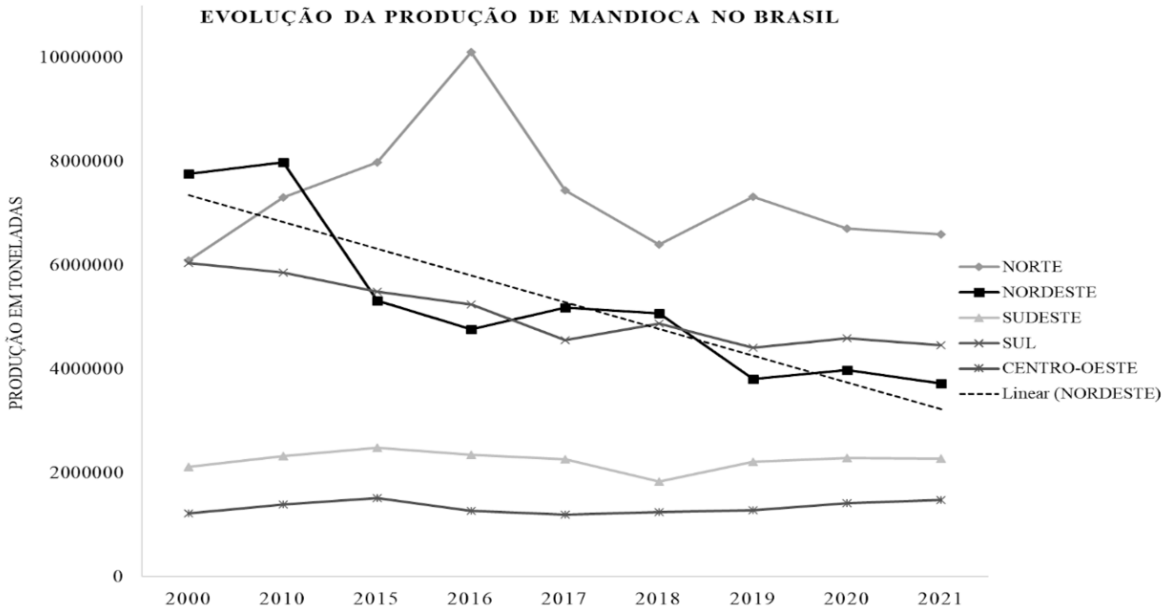
1. A cultura da mandioca

A mandioca (*Manihot esculenta* Crantz.), família Euphorbiaceae, é uma planta nativa do Brasil com provável centro de origem ao sul da bacia do rio Amazonas. A domesticação da cultura teve como ancestral a subespécie selvagem *M. esculenta* ssp. *flabellifolia* (MARTINS *et al*., 2020; OLSEN, 2004).

A cultura da mandioca é uma importante fonte de subsistência nas regiões tropicais (MCCALLUM *et al*., 2017). Segundo dados da FAO, no ano de 2020 a produção mundial de mandioca foi de 302.662.494 toneladas. Os principais países produtores em 2020 foram a Nigéria, República Dominicana e Tailândia. O Brasil foi o sexto maior produtor, com uma produção de 18.205.120 toneladas.

A região Nordeste do Brasil teve ao longo dos anos uma queda na produção, quando comparados os anos de 2000 (7.755.879 t) e 2021 (3.719.184 t), representando aproximadamente 48% de queda segundo dados do IBGE de 2022 (Figura 1). Dentro desse cenário, o estado de Pernambuco foi responsável pela produção de 435.695 toneladas em 2021, com área plantada de 43.584 ha (IBGE, 2022). Embora o cultivo no estado seja pouco tecnificado e a área de plantio tenha decaído nas últimas décadas, a cultura tem uma grande importância na alimentação e é alternativa socioeconômica para pequenas propriedades na região (CUECA; MANDARINO, 2006).

Figura 1. Gráfico da evolução da produção de mandioca no Brasil nas últimas décadas, segundo dados do IBGE.



Fonte: IBGE, 2022.

A mandioca, por ser uma cultura mundialmente importante, está na lista de espécies cultivadas do sistema multilateral do Tratado Internacional sobre Recursos Filogenéticos para a Alimentação e a Agricultura. O Tratado tem como objetivo principal a manutenção dos recursos genéticos para agricultura, garantindo a segurança alimentar (BRASIL, 2008). Visando a manutenção dos recursos genéticos no Brasil, diversos órgãos de Pesquisa mantêm Banco Ativo de Germoplasma (BAG) no país. Atualmente, o Ministério da Agricultura, Pecuária e Abastecimento do Brasil possui atualmente o registro oficial de 104 cultivares de mandioca mantidas por órgãos de pesquisa do País, entre eles: a Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), o Instituto de Desenvolvimento Rural do Paraná (IAPAR-EMATER), o Instituto Agronômico (IAC), o Instituto Agronômico de Pernambuco (IPA), a Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina (EPAGRI), a Secretaria da Agricultura, Pecuária e Desenvolvimento Rural (DDPA), a Universidade de Brasília (UNB), o Instituto Centro de Ensino Tecnológico (CENTEC- CEPLAC) e a Comissão Executiva do Plano da Lavoura Cacaueira (CEPLAC) (MAPA, 2022).

As cultivares de mandioca variam de acordo com o teor de ácido cianídrico (HCN) acumulado nas raízes. As com baixo teor são popularmente chamadas de aipim, macaxeira ou

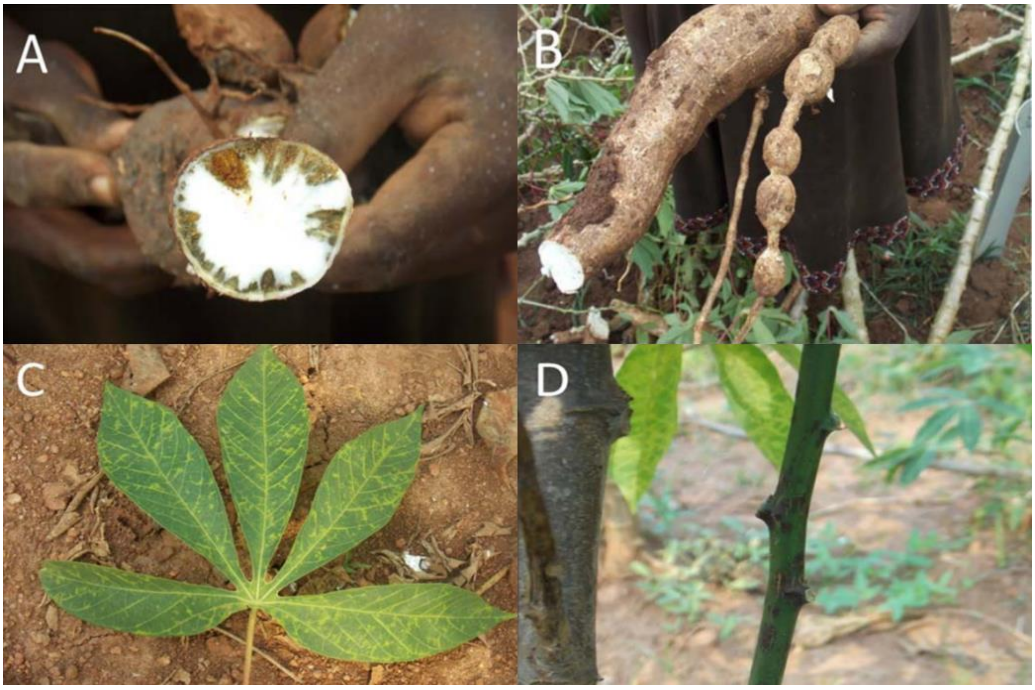
mandioca mansa e as plantas com raízes que contém um alto teor de HCN são conhecidas como amargas ou mandioca bravas. De acordo com essa classificação, a mandioca pode ser utilizada como mandioca de mesa, comercializada geralmente *in natura* ou mandioca para indústria, que é processada em diversos subprodutos, tendo um papel importante na geração de emprego, especialmente em regiões com baixa renda *per capita* (MATTOS; CARDOSO, 2003).

O plantio das manivas deve ocorrer em condições ideais de umidade para que haja a brotação, mas não favoreça o desenvolvimento de patógenos radiculares. A cultura pode se adaptar e desenvolver em uma faixa de temperatura que varia de 16 a 38ºC. A principal prática cultural indicada para a cultura é o manejo de plantas daninhas, o que pode representar 35% dos custos de produção (MATTOS; CARDOSO, 2003).

2. Principais viroses da cultura da mandioca

As doenças da cultura da mandioca representam um desafio e podem impactar amplamente a sua produção, contribuindo para a redução do total produzido (MCCALLUM *et al*., 2017). Os vírus estão entre os principais patógenos. Os mais importantes no cenário mundial atualmente são o agente causal da doença da estria marrom (*Cassava Brown Streak Virus Disease*, CBSD) e os vírus associados a doença do mosaico da mandioca (*Cassava Mosaic Disease*, CMD) (Figuras 2 e 3). Essas doenças são consideradas os mais devastadores na África (WINTER *et al*., 2010; ZINGA *et al*., 2013), e podem ser responsáveis por diminuir a produção de raízes em 70% (HILLOCKS *et al*., 2008), podendo chegar a 80% de redução com a ocorrência de infecções mistas (OWOR *et al.*, 2004).

Figura 2. Sintomas da estria marrom da mandioca.



A) Necrose nas raízes, B) Constrições nas raízes, C) clorose foliar e (D) lesões escuras nas hastes. Fonte: Retirado de TOMLINSON et al. (2018).

Figura 3. Sintomas do mosaico da mandioca.

A, B, C e D) Sintomas de mosaico, C e D) Folhas retorcidas, E) desfolhamento. Fonte: Retirado de MULENGA et al. (2016).

A forma de propagação vegetativa da mandioca pode contribuir para a disseminação e introdução de novos patógenos nas regiões produtoras, agravando a atual situação sanitária da cultura (ENI *et al.*, 2020; ZINGA *et al*., 2013). Com o objetivo de evitar a entrada de novos virus no Brasil, em 2019 foram publicados subsídios para inclusão do complexo de todas as espécies da família *Geminiviridae* que causam o mosaico da mandioca africano na lista de pragas quarentenárias A1 do Brasil, ressaltando a importância do monitoramento dos vírus na cultura, pois as espécies presentes na África representam alto risco para a cadeia produtiva de mandioca no país (ANDRADE; LARANJEIRA 2019).

3. Principais vírus da cultura da mandioca no Brasil

As doenças mais importantes induzidas na mandioca relatadas no Brasil são o mosaico-das-nervuras, causada pelo potexvírus CsVMV (cassava vein mosaic virus), o mosaico-comum, causado pelo cavemovírus CsCMV (cassava common mosaic virus) e possivelmente os vírus que podem estar associados a doença das raízes denominada de "couro de sapo" cuja etiologia ainda não foi elucidada (ANDRADE; LARANJEIRA 2019). No início da década de 1970, a doença couro de sapo da mandioca (*Cassava Frogskin Disease*, CFSD) foi descrita na Colômbia causando grandes perdas (PINEDA *et al.*, 1983) e tem sido associada a sintomas radiculares em plantas de mandioca no Brasil, Costa Rica, Panamá, Peru e Venezuela. Além disso, alguns autores já associaram a CFSD à fitoplasmas (cepas do grupo 16SrIII) (ALVAREZ *et al.*, 2009). No entanto, estudos anteriores mostraram que a diversidade viral nas amostras de mandioca pode estar sendo subestimada e a presença de infecções virais mistas na cultura ainda precisa ser estudada (CARVAJAL-YEPES *et al*., 2014; ZINGA *et al*., 2013).

Além do cavemovírus CsVMV e do potexvírus CsCMV, já foram relatados no Brasil: o cassava American latente (CsALV, *Nepovirus*) (WALTER *et al.*, 1989), cassava polero-like (CsPLV, *Polerovirus*), cassava new alphaflexivirus (CsNAV, *Potexvirus*) e o cassava Torrado-like (CsTLV, *Torradovirus*) (OLIVEIRA *et al*., 2020). Desses, apenas o CsCMV, CsVMV, CsTLV e CsPLV foram detectados por técnicas moleculares (OLIVEIRA *et al*., 2020; COLARICCIO *et al*., 2020).

3.1. Gênero *Potexvirus*

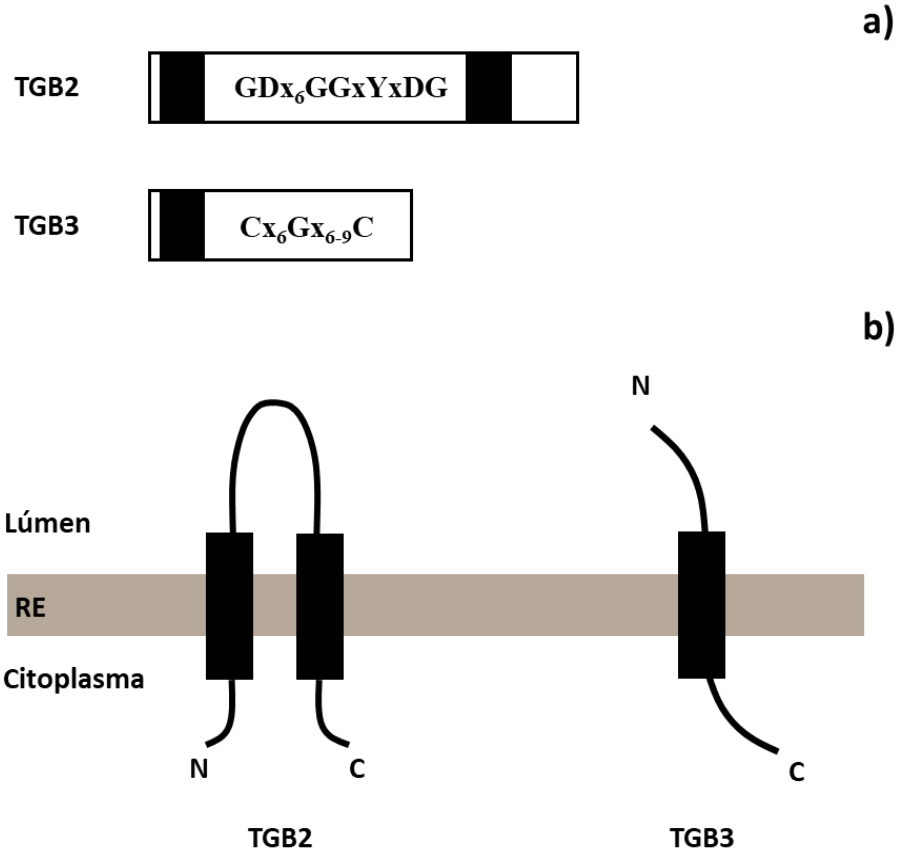
Os potexvírus pertencem à família *Alphaflexiviridae*. O gênero *Potexvirus* possui atualmente 49 espécies aprovadas no Comitê Internacional de Taxonomia de vírus (International Committee on Taxonomy of Viruses, ICTV). A espécie viral que deu nome ao

gênero foi o *Potato virus X* (ICTV, 2022). Nesse gênero três vírus já foram relatados infectando mandioca e são reconhecidas pelo ICTV: CsCMV (COSTA, 1940; KITAJIMA *et al*., 1965), CsVX (CHAPARRO-MARTÍNEZ; TRUJILLO-PINTO, 2001); e cassava colombian symptomless virus (CsCSV) (descritos anteriormente como cassava caribbean mosaic virus, CsCAMV; e cassava new alphaflexivirus, CsNAV) (ALABI; KUMAR; NAIDU, 2011; CARVAJAL-YEPES *et al*., 2014). Doenças virais como a doença do mosaico da mandioca causada pelo potexvírus CsCMV podem ser responsáveis por perdas de até 30% na produção brasileira da cultura (VENTURINE *et al*., 2016). Os potexvírus são conhecidos por serem transmitidos mecanicamente e a disseminação através de estacas de mandioca infectadas pode promover a propagação viral (LOZANO *et al*., 2017).

A partícula viral é filamentosa flexuosa com simetria helicoidal, diâmetro de 13 nm e comprimento que varia de 470 a 580 nm, dependendo da espécie. O genoma viral codifica até cinco ORFs, e é composto por uma molécula de RNA linear positiva com tamanho variando de 5,9 a 7.0 kb. A molécula de RNA genômico possui uma estrutura de *cap* (quepe) no terminal 5' e uma cauda poli-A na extremidade 3' (SONENBERG *et al*., 1978; ATABEKOV *et al*., 2007; KREUZE *et al*., 2020).

A proteína codificada pela ORF1 nos potexvírus contém os domínios de RNA polimerase dependente de RNA (RdRp), metiltransferase (Met), e de helicase (Hel). As três ORFs responsáveis por codificar o gene triplo bloco (TGB, *Triple Gene Block*), estão em quadros de leitura diferentes e são expressos a partir de RNAs subgenômicos. A proteína TGB2 dos potexvírus possui dois domínios transmembranares, que auxiliam na integração dessa proteína às membranas do retículo endoplasmático (RE), formando uma estrutura em forma de U. Nesta estrutura os terminais N e C ficam localizados no citoplasma do RE. A proteína TGB3 dos potexvírus possui apenas um domínio transmembranar, onde a região N-terminal com carga negativa é exposta ao lúmen do RE e a C-terminal é exposta ao citoplasma (Figura 4). No modelo de transporte dos potexvírus, as proteínas do TGB e CP são necessárias para o movimento célula-à-célula (MOROZOV; SOLOVYEV, 2003; PARK *et al*., 2014).

Figura 4. Representação das proteínas TGB2 e TGB3 encontradas nos potexvírus, com as regiões conservadas e organização molecular de TGB2 e TGB3.



A) Regiões conservadas das proteínas, B) Topologia putativa das moléculas TGB2 e TGB3 na membrana do retículo endoplasmático (RE). Os retângulos pretos representam os domínios transmembrana. Os motivos conservados das proteínas estão representados. Fonte: Adaptado de MOROZOV; SOLOVYEV, 2003.

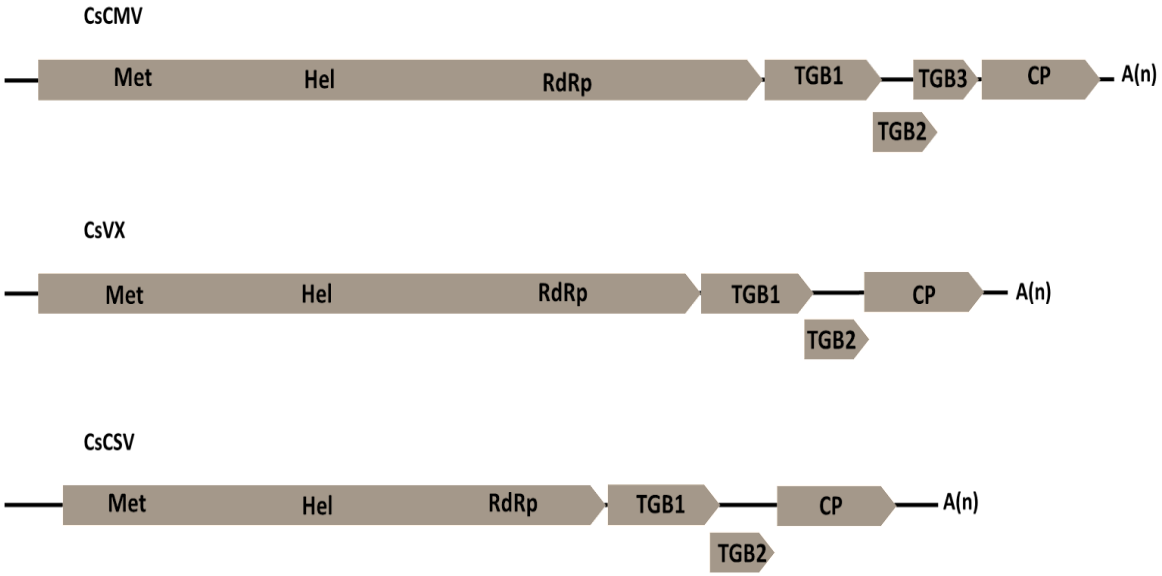
O vírus do mosaico comum da mandioca (CsCMV) foi relatado pela primeira vez infectando plantas de *M. esculenta* em 1940 no Brasil (COSTA, 1940). O vírus também foi relatado infectando plantas de *Cnidoscolus chayamansa* no ano de 1998 em Tuvalu e em 2015 na Venezuela (JONES *et al*. 1998; MEJÍAS *et al.*, 2015) e novamente, recentemente, entre os anos de 2018 e 2021, o CsCMV foi relatado infectando plantas de mandioca na América do Sul (COLLAVINO *et al.*, 2021; ZANINI *et al.*, 2018) e China (TUO *et al*., 2020; YU *et al*., 2020). Infecções mistas são comuns entre os potexvírus, porém o único capaz de causar doença em infecções simples em mandioca é o CsCMV (LOZANO *et al.*, 2017). As perdas na cultura podem chegar a 30% no Brasil (VENTURINE *et al*., 2016).

O RNA genômico do CsCMV é de fita simples positivo e possui cinco ORFs. A ORF1 é responsável por codificar a proteína associada à replicação (RdRp). As ORFs 2, 3 e 4

codificam as proteínas do bloco triplo (TGBs), e são envolvidas no movimento célula-a-célula. Por fim, a ORF5 codifica a capa proteica (CP), que também está envolvida no movimento célula-a-célula. A organização genômica do CsVX e CsCSV difere do CsCMV, pois não possuem a ORF responsável por codificar o TGB3 (Figura 5) (PARK; JEONG; KIM, 2014; KREUZE *et al*., 2020).

O CsVX foi reportado pela primeira vez em 1985 infectando mandioca na América do Sul (HARRINSON *et al*., 1986, LENNON *et al*., 1985), e já foi encontrado em infecções mistas com CsCSV na Colômbia em plantas de campo apresentando sintomas de couro de sapo (HARRISON *et al*., 1986, LENNON *et al*., 1986, NOLT *et al.*, 1992). Segundo Chaparro-Martínez e TrujillPinto (2001), plantas de mandioca cultivadas infectadas no campo com CsVX não expressaram sintomas. No entanto, a transmissão de CsVX através de estacas infectadas é 100% efetiva (CHAPARRO-MARTÍNEZ; TRUJILLO-PINTO, 2001). Lozano et al. (2017) mostraram que plantas de mandioca infectadas com CsVX e CsCSV não expressaram sintomas em experimentos de transmissão mecânica (LOZANO *et al*., 2017). No entanto, plantas de *Chenopodium quinoa* Willd. e de *Nicotiana benthamiana* Domin. infectadas com CsVX mostraram sintomas virais após inoculação (LENNON *et al.*, 1986; LOZANO *et al*., 2017).

Figura 5. Organização genômica do cassava common mosaic virus (GenBank NC_001658.1), cassava virus X (GenBank NC_034375) e cassava colombian symptomless virus (GenBank KC505252).



As extremidades 5’e 3’ não traduzidas (UTR) são representadas por um segmento de reta. As ORFs que codificam as proteínas da replicase RdRp (*RNA-dependent RNA polymerase*), TGB1, TGB2, TGB3 (*Triple Gene Block e* CP (capa proteica) são representadas em cinza. A região da cauda poli-A foi representada pela notação A(n).

3.2. Gênero *Cavemovirus*

Os cavemovírus pertencem à família *Caulimoviridae*. O CsVMV foi descoberto em 1966 na América do Sul, Brasil e deu origem ao nome no gênero. Os cavemovírus possuem uma partícula viral icosaédrica, não envelopadas com cerca de 50 nm de diâmetro e seu material genético é composto por DNA de fita dupla circular (ICTV, 2022; TEYCHENEY *et al.*, 2020; KITAJIMA; COSTA, 1966). Além do CsVMV, mais duas espécies foram aprovadas pelo ICTV: o *Epiphyllum virus 4* (MN153807, EpV4), aceito em 2020, e o *Sweet potato collusive virus*, aceito em 2011(NC_015328, SPCV) (ICTV, 2022; ZHENG *et al.*, 2020; CUELLAR *et al.*, 2011; de SOUZA; CUELLAR, 2011).

Os cavemovírus codificam proteínas virais a partir de quatro a cinco ORFs (ZHENG *et al.*, 2020; CALVERT OSPINA; SHEPHERD, 1995). A ORF1 codifica uma poliproteína que contém a capa proteica e domínios da proteína de movimento. A ORF2 codifica uma proteína de função desconhecida. A ORF3 codifica uma poliproteína com domínios de protease aspártica, transcriptase reversa (RT) e de ribonuclease RNaseH. A ORF4 codifica o transativador de proteína de tradução. A ORF 5 codifica uma proteína de função desconhecida (CALVERT *et al.*, 1995).

Os vírus desse gênero são transmitidos por propagação vegetativa, inoculação mecânica e afídeos (ICTV, 2022; ZHENG *et al.*, 2020), os sintomas expressos em infecção simples são de clareamento de nervuras (CUELLAR *et al.*, 2011), mosaicos, distorção foliar e epinastia em folhas jovens (CALVERT *et al*., 1995).

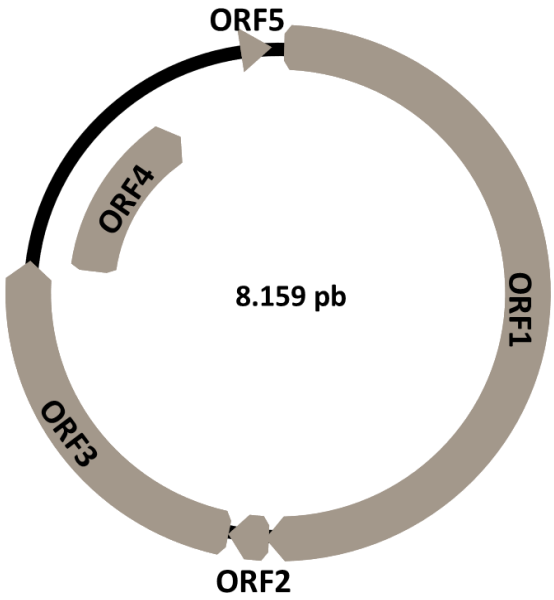
O primeiro relato do cavemovírus que causa mosaico das nervuras da mandioca (cassava vein mosaic virus, CsVMV) foi em 1966 na América do Sul, Brasil. As partículas virais de CsVMV têm 50–60 nm de diâmetro (KITAJIMA; COSTA, 1966).

O genoma viral possui uma região intergênica, na qual estão localizados o promotor de RNA pré-genômico, um sinal de poliadenilação do RNA e sítio de ligação do iniciador de cadeia secundária (KING et al., 2012). A ORF1 do seu genoma codifica uma poliproteína que contém a capa proteica e domínios da proteína de movimento. A ORF2 codifica uma proteína de função desconhecida. A ORF3 codifica uma poliproteína com a protease aspártica, os domínios da transcriptase reversa (RT) e a ribonuclease RNaseH. A ORF4 codifica o

transativador de proteína de tradução. A ORF 5 codifica uma proteína de função desconhecida (CALVERT *et al.*, 1995) (Figura 6).

O CsVMV é transmitido através de estacas infectadas utilizadas para propagação da cultura (TEYCHENEY *et al.*, 2020), que ao germinarem formam plantas atrofiadas. Infecções por CsVMV em plantas de mandioca podem causar doenças e expressar sintomas de clorose nas nervuras que podem coalescer e formar mosaicos, induzir a distorção foliar e epinastia em folhas jovens (CALVERT *et al.*, 1995). Esses sintomas são transitórios e as folhas seguintes podem não apresentar sintomas (CALVERT *et al.*, 1995).

Figura 6. Organização genômica do cassava vein mosaic virus (GenBank NC_001648.1).



A ORF1 codifica uma poliproteína que contém a capa proteica e domínios da proteína de movimento. A ORF2 codifica uma proteína de função desconhecida. A ORF3 codifica uma poliproteína com a protease aspártica, os domínios da transcriptase reversa (RT) e a ribonuclease RNaseH. A ORF4 codifica o transativador de proteína de tradução. A ORF 5 codifica uma proteína de função desconhecida (CALVERT; OSPINA; SHEPHERD, 1995).

3.3. Gênero *Torradovirus*

O gênero *Torradovirus*, família *Secoviridae*, possui apenas seis espécies aceitas pelo ICTV: *Carrot torradovirus 1* (NC_025479/NC_025480, CaTV1), *Lettuce necrotic leaf curl virus* (NC_035214/NC_035219, LNLCV), *Motherwort yellow mottle virus* (NC_035218/ NC_035220, MYMoV), *Squash chlorotic leaf spot virus* (NC_035221/NC_035215, SCLSV), *Tomato marchitez virus* (NC_010987/ NC_010988, ToMarV), *Tomato torrado virus* (NC_009013/ NC_009032, ToTV) (ICTV, 2022). O vírus que deu origem ao gênero

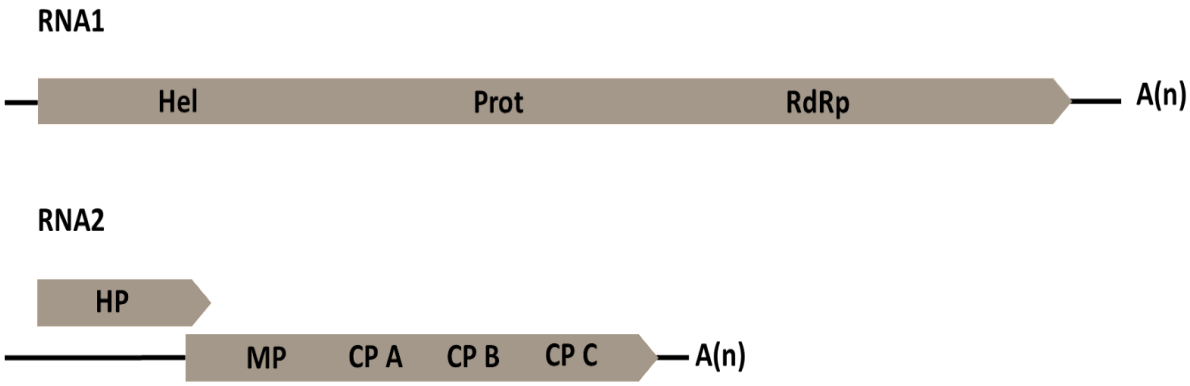
Torradovirus, tomato torrado virus, foi relatado em 2007 (VERBEEK *et al.*, 2007), e desde então novas espécies do gênero foram relatadas em plantas de diferentes famílias botânicas, como cenoura, alface, agripalma, abóbora e mandioca (WINTERMANTEL *et al*., 2018; LECOQ *et al*., 2016; SEO *et al.*, 2015; van der VLUGT *et al.*, 2015, ADAMS *et al.*, 2014 VERBEEK *et al.*, 2014; BATUMAN *et al*., 2010; VERBEEK *et al.*, 2010; VERBEEK *et al.*, 2008; VERBEEK *et al*., 2007, CARVAJAL-YEPES *et al.*, 2014).

Os vírus desse gênero têm partículas isométricas não envelopadas e contêm um genoma composto por dois RNAs de fita de senso positiva de fita única. O RNA1 e RNA2 tem tamanho variado de 6911-7802 nucleotídeos (nt) e 5695-4701 nt, respectivamente, para espécies relatadas e aceitas pelo ICTV. O RNA1 codifica uma poliproteína de 2151-2224 aminoácidos (aa), que é processada em um cofator de protease, uma helicase, uma protease semelhante a 3C e uma RNA polimerase dependente de RNA. O RNA2 codifica uma proteína hipotética (185-212 aa) e uma poliproteína (1168-1223 aa), que é processada em uma proteína de movimento (MP) e três proteínas de revestimento (CP-A, CP-B e CP-C). Os RNAs genômicos dos torradovírus geralmente têm uma 3'-UTR longa e uma cauda poli(A) (THOMPSON *et al*., 2017; FERRIOL *et al*., 2016; van der VLUGT *et al*., 2015).

A transmissão dos torradovírus pode ocorrer de forma semi-persistente por espécies de mosca branca, tais como, *Bemisia tabaci* Genn., *Trialeurodes vaporariorum* Westwood e *Trialeurodes abutilonea* Haldeman (LECOQ *et al*., 2016; VERBEEK *et al.*, 2014) e por espécies de pulgões, *Cavariella aegopodii* Scopoli. e *Myzus persicae* Sulzer. (ROZADO-AGUIRRE *et al.*, 2016).

Os sintomas típicos de torradoviroses são manchas necróticas com halos amarelos em tomateiro (van der VLUGT *et al.*, 2015), mosqueados amarelos, mosaicos e nanismo em *Leonurus sibiricus* L. (SEO *et al.*, 2014). Além disso, o CsTLV foi detectado em amostras com sintomas de couro de sapo em mandioca (CARVAJAL-YEPES *et al*., 2014; OLIVEIRA *et al*., 2020), mas não foi comprovada relação com a doença. O CsTLV foi identificado pela primeira vez na mandioca na Colômbia em 2014 (CARVAJAL-YEPES *et al*., 2014). O genoma do CsTLV é bipartido, semelhante aos genomas do gênero. Na poliproteína expressa pelo RNA1 é encontrada a proteína responsável pela replicação viral, RdRp. No RNA2 está presente uma proteína hipotética na ORF1 e é codificada a proteína de movimento, onde está presente o motivo proteico LxxPxL e as três moléculas da capa proteica (CP-A, CP-B e CP-C) na poliproteína da ORF2 (Figura 7) (van der VLUGT *et al.*, 2015).

Figura 7. Organização genômica do cassava Torrado-like virus (GenBank RNA1-MF449522.1 e RNA2-MF449523.1).

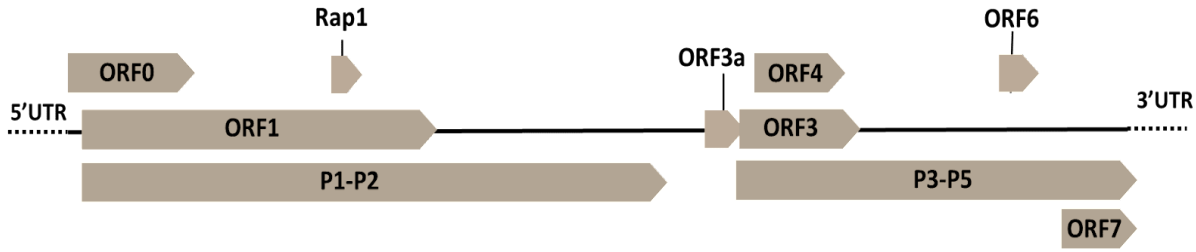


As ORFs que codificam as proteínas da replicase (RdRp, *RNA-dependent RNA polymerase*) no RNA1, proteína hipotética (HP), proteína de movimento (MP), capa proteica (CP-A, CP-B e CP-C) no RNA2 são representadas em cinza. Os motivos da helicase (Hel) e protease (Prot) estão indicados na ORF1. As extremidades 5'e 3' não traduzidas (UTR) são representadas por um segmento de reta e a região da cauda poli(A) foi representada pela notação A(n)

3.4 Gênero *Polerovirus*

As partículas de polerovírus são icosaédricas não-envelopadas (BYRNE *et al.*, 2019). O RNA genômico de sentido positivo é policistrônico e codifica até 10 ORFs (DELFOSSE *et al.*, 2021; JEEVALATHA *et al.*, 2013; JAAG *et al.*, 2003). A tradução da ORF0, ORF1 e ORF2 ocorre a partir do RNA genômico. Na tradução da ORF0, os ribossomos usam um mecanismo de varredura com vazamento (*leaky scanning*) para iniciar a leitura da ORF1 e o mecanismo de mudança de quadro ribossomal -1 (-1 *ribosomal frameshift*) acontece na ORF1 para a leitura da ORF2. A ORF2 codifica a RNA polimerase dependente de RNA. O subgenômico RNA1 é usado para a tradução das ORFs 3a, 3, 4 e 5. Enquanto um RNA2 subgenômico é usado para a tradução da ORF6 e ORF7 e um RNA3 subgenômico apenas para a tradução da ORF7. A ORF3 e ORF5 são fusionadas e a expressão da ORF5 ocorre por leitura ribossomal (*readthrough*) a partir do códon de parada da ORF3 (DELFOSSE *et al.*, 2021; SÕMERA *et al.*, 2021; SMIRNOVA *et al.*, 2015; JAAG *et al.*, 2003). A ORF8 foi encontrada em potato leafroll virus (PLRV), responsável por codificar a proteína Rap1, que é associada a replicação viral (Figura 8) (JEEVALATHA *et al.*, 2013).

Figura 8. Organização genômica do vírus do enrolamento da batata (potato leafroll virus, PLRV).



As ORFs são representadas em cinza e as extremidades 5'e 3' não traduzidas (UTR) são representadas por linhas pontilhadas. Fonte: Adaptado de SÕMERA *et al*., 2021.

A infecção viral por polerovírus em plantas ocorre no floema e alterações no proteoma deste tecido parecem contribuir para a infecção viral nas plantas e na transmissão por pulgões (CILIA *et al.*, 2012; GRAY *et al*., 2003). A transmissão também pode ser feita pela mosca branca (*Bemisia tabaci*), exemplo encontrado nas transmissões de pepper whitefly-borne vein yellows virus (GHOSH *et al.*, 2019) e de cucurbit aphid-borne yellows virus (COSTA *et al*., 2020).

Os sintomas expressos em plantas infectadas por polerovírus incluem amarelecimento, avermelhamento, enrolamento, folhas pequenas, descoloração da nervura principal da folha, aparência espessa e quebradiça, encurtamento e nanismo nos entrenós (ABRAHAM *et al*., 2022; ZHAO *et al*., 2021; DISTÉFANO *et al.*, 2010).

O polerovírus que infecta mandioca, de acordo com o ICTV (2022), cassava polero-like virus foi encontrado em infecções mistas em plantas de mandioca que expressam sintomas de couro de sapo em raízes no Brasil (OLIVEIRA *et al*., 2020) e na Colômbia (CARVAJAL-YEPES *et al*., 2014). Porém não há estudos complementares associando o vírus à doença.

4. Manejo das viroses na cultura da mandioca

O manejo das viroses da cultura da mandioca deve iniciar no planejamento do plantio, com a utilização de mudas sadias, uma vez que a dispersão viral por meio de material propagativo infectado pode contribuir para introdução de vírus em áreas livres. Além disso, deve-se dar preferência ao plantio de genótipos tolerantes ou resistentes aos vírus, pois com o estabelecimento de viroses no campo, seria possível minimizar os danos através do controle genético (ANDRADE *et al*., 2020; ANDRADE; LARANJEIRA 2019). Outra estratégia

importante é a escolha da área de plantio longe de possíveis fontes de inóculo de plantações já infectadas (LEGG *et al*., 2014).

As técnicas de termoterapia e cultura de tecidos a partir de meristemas têm sido utilizadas para a eliminação de vírus em germoplasmas de mandioca (KIDULILE *et al*., 2018; ACHEREMU *et al*., 2015). Porém, é preciso que se desenvolva métodos de detecção rápidos e eficientes, que contribuam com a diagnose viral e o estabelecimento de medidas quarentenárias, para garantir que as mudas produzidas após a limpeza clonal com essas técnicas estejam realmente livres de vírus (ANDRADE *et al*., 2020).

Medidas fitossanitárias como a desinfestação de ferramentas agrícolas utilizadas no manejo da cultura, principalmente as de corte, pode impedir a transmissão viral entre plantas de uma mesma área. As medidas de controle que também podem ser tomadas após o aparecimento de viroses no campo incluem a eliminação de plantas infectadas 'roguing'; controle de insetos vetores, quando for viável, e controle de plantas que possam ser hospedeiras alternativas dos vírus (LEGG *et al*., 2014).

5. Sequenciamento de alto rendimento (*High Throughput Sequencing*-HTS) e sua aplicação da detecção e estudos sobre a diversidade viral

Os vírus são importantes agentes que podem modular a evolução de micro-organismos e têm a capacidade de reprogramar e interferir em rotas metabólicas de seus hospedeiros (HURWITZ; HALLAM; SULLIVAN, 2013; ROHWER; THURBER, 2009; MARTÍNEZ-GARCÍA *et al*., 2014). Porém, estudos indicam que a diversidade viral e seu papel em ecossistemas ainda precisa ser elucidada (ROUX *et al*., 2015). Estima-se que nos oceanos 10 milhões de partículas por mililitro são virais, o que os torna as entidades biológicas em maior número nos oceanos (BREITBART *et al*., 2007). Análises realizadas em bancos públicos de metadados mostraram que de um total de 3.376.880, 59,38% das execuções continham prováveis sequências de polimerase viral do tipo RdRp (EDGAR *et al.*, 2022). A análise de metadados em metagenômica pode contribuir para compreensão da dinâmica temporal e espacial de comunidades e permitirá a caracterização e estudo dos vírus presentes no sistema, o que ajuda a entender o impacto dos vírus no ecossistema (BREITBART *et al*., 2007; MIZUNO *et al*., 2013). Dentro deste contexto, diversas ferramentas foram lançadas e aprimoradas para detecção de sequências virais em amostras de metadados (GUO *et al*., 2021; ROUX *et al*., 2015), especialmente após a pandemia mundial causada pelo vírus da síndrome respiratória aguda grave (S*evere acute respiratory syndrome coronavirus 2*; SARS-CoV-2),

com o lançamento da plataforma *Serratus* (https://serratus.io/) (EDGAR *et al.*, 2022). Além disso, foram desenvolvidos *pipelines* práticos e específicos para descoberta de novos vírus de plantas com a análise de metadados (BLAWID *et al*., 2017), assim como para a análise de diferentes tipos de amostras (PLYUSNIN *et al.*, 2020).

O sequenciamento de alto rendimento (HTS) é uma importante ferramenta para elucidar a diversidade viral em plantas (MALIOGKA ***et al.*, 2018**; PECMAN *et al*., 2017; BARBA; CZOSNEK; HADIDI, 2014), além de contribuir para a detecção e monitoramento de vírus em material propagativo (REIS *et al*., 2020; VILLAMOR *et al*., 2019). Diversos estudos foram realizados com HTS para detecção de vírus de plantas e entendimento da diversidade viral em amostras vegetais, mostrando a importância da metagenômica para o desenvolvimento da virologia de plantas (SANTOS *et al*., 2022; MALIOGKA *et al.*, 2018; ROTT *et al.*, 2018; VILLAMOR *et al.*, 2016; ESPACH; MAREE; BURGER, 2012). Desde 2017 o ICTV tem feito esforços para que haja a aceitação de propostas de novas espécies descobertas a partir de dados de HTS, uma vez que o número de depósitos de sequências virais tem crescido significativamente e consequentemente gera um grande número de informação para o desenvolvimento da taxonomia viral (SIMMONDS *et al*., 2017).

Esse tipo de tecnologia é capaz de sequenciar em paralelo milhões ou até mesmo bilhões de cadeias de DNA, possibilitando a detecção de variantes virais e a caracterização da estrutura da população viral de *quasi-species* (NATURE PORTFOLIO, 2019). É crescente a evolução na qualidade e no rendimento do HTS nos últimos anos. Além disso, a tecnologia tem se tornado mais acessível, impulsionando o surgimento de novas tecnologias de processamento de dados. O HTS pode ser utilizado em estudos integrados com sequenciamento híbrido; de leitura longa e leitura curta, sequenciamento clínico de rotina, monitoramento de patógenos em tempo real e estudos a nível de população (GOODWIN; McPHERSON; McCOMBIE, 2016).

Os sequenciamentos de última geração, combinado aos métodos tradicionais em virologia de plantas podem ser usados para estudar as interações presentes em um ecossistema. A metagenômica tem sido responsável por facilitar a identificação de vírus, tornando-a mais rápida. Sequenciadores portáteis, como o MinION (*Oxford Nanopore Technologies*), já estão sendo utilizados em campo para o sequenciamento de genomas virais. Estima-se que na próxima década, esses métodos ajudarão na compreensão das interações entre vírus-hospedeiro-vetor, assim como na forma de entendimento da disseminação dos vírus, no estudo do surgimento de novos hospedeiros e no entendimento da introdução de novos fitopatógenos através de mudas infectadas (LEFEUVRE *et al*., 2019). Adicionalmente, os eventos de

recombinação nos vírus podem ser detectados com o sequenciamento HTS (BÖMER *et al*., 2018).

Diante do exposto, esta pesquisa visou o estudo da diversidade viral em amostras de (meta)transcriptomas de mandioca, entender a relação filogenética e a variabilidade dos genomas montados, assim como sugerir medidas de controle para prevenir a dispersão e trânsito de material vegetal propagativo infectado, através da detecção e monitoramento com sequenciamento de alto rendimento.

REFERÊNCIAS BIBLIOGRÁFICAS

ABRAHAM, A.; VETTEN, H. J. Chickpea chlorotic stunt virus: a threat to cool-season food legumes. **Archives of Virology**, v. 167, p. 21-30, 2022.

ACHEREMU, K.; AKROMAN, R.; YUSSSF JNR, I.; SANTO, K. G. Effect of thermotherapy in the elimination of viruses on four (4) mosaic diseased cassava cultivars. **Journal of Biology, Agriculture and Healthcare**, v. 5, n. 22, p. 1-15, 2015.

ADAMS, I. P.; SKELTON, A.; MACARTHUR, R.; HODGES, T.; HINDS, H.; FLINT, L.; NATH, P. D.; BOONHAM, N.; FOX, A. Carrot yellow leaf virus is associated with carrot internal necrosis. **PLoS One**, v. 9, n. 11, p. e109125, 2014.

ALABI, O. J.; KUMAR, P. L.; AND NAIDU, R. A. 2011. **Cassava mosaic disease:** A curse to food security in Sub-Saharan Africa. St. Paul, 2001. APSnet Features. doi:10.1094/APSnetFeature-2011-0701. Disponível em: <https://www.apsnet.org/edcenter/apsnetfeatures/Pages/cassava.aspx>. Acesso em: 04 fev. 2022.

ALVAREZ, E.; MEJÍA, J. F.; LLANO, G. A.; LOKE, J. B.; CALARI, A.; DUDUK, B.; BERTACCINI, A. Characterization of a phytoplasma associated with frogskin disease in cassava. **Plant Disease**, v. 93, n. 11, p. 1139-1145, 2009.

ANDRADE, E. C.; LARANJEIRA, F. F. **African cassava mosaic virus (ACMV) e a Doença do Mosaico da Mandioca (Cassava Mosaic Disease, CMD):** Subsídios para inclusão de novas espécies de begomovírus causadoras do CMD na lista de pragas quarentenárias A1 do Brasil. Cruz das Almas: Embrapa Mandioca e Fruticultura, 2019. 22 p. (Embrapa Mandioca e Fruticultura. Documentos, 240).

ANDRADE, E. C.; MEISSNER FILHO, P. E.; OLIVEIRA, S. A. S.; BARBOSA, F. F. L. **Subsídios para plano de contingência da "Doença das Estrias Marrons (CBSD)" e do**

"Mosaico Africano (CMD)" da mandioca. Cruz das Almas: Embrapa Mandioca e Fruticultura, 2020. 32 p. (Embrapa Mandioca e Fruticultura. Documentos, 241).

ATABEKOV, J.; DOBROV, E.; KARPOVA, O.; RODIONOVA, N. Potato virus X: structure, disassembly and reconstitution. **Molecular Plant Pathology**, v. 8, n. 5, p. 667-675, 2007.

BARBA, M.; CZOSNEK, H.; HADIDI, A. Historical perspective, development and applications of next-generation sequencing in plant virology. **Viruses**, v. 6, n. 1, p. 106-136, 2014.

BATUMAN, O.; KUO, Y. W.; PALMIERI, M.; ROJAS, M. R.; GILBERTSON, R. L. Tomato chocolate spot virus, a member of a new torradovirus species that causes a necrosis-associated disease of tomato in Guatemala. **Archives of Virology**, v. 155, n. 6, p. 857-869, 2010.

BÖMER, M.; RATHNAYAKE, A. L.; VISENDI, P.; SILVA, G.; SEAL, S. E. 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**, v. 163, n. 2, p. 533-538, 2018.

BRASIL. Decreto nº 6.476 de 05 de junho de 2008. Promulga o Tratado internacional sobre recursos fitogenéticos para a alimentação e a agricultura, aprovado em Roma, em 3 de novembro de 2001, e assinado pelo Brasil em 10 de junho de 2002. **Diário Oficial da União**: Secão I, Brasília, n. 107, p. 8-14, 2008.

BREITBART, M.; THOMPSON, L.R.; SUTTLE, C.A.; SULLIVAN. M.B. Exploring the vast diversity of marine viruses. **Oceanography**, v. 20, n. 2, p.135–139, 2007.

BYRNE, M. J.; STEELE, J. F. C.; HESKETH, E. L.; WALDEN, M.; THOMPSON, R. F.; LOMONOSSOFF, G. P.; RANSON, N. A. Combining transient expression and Cryo-EM to obtain high resolution structures of luteovirid particles. **Structure,** Londres, v. 27, n. 12, p. 1761-1770, 2019.

CALVERT, L. A.; OSPINA, M. D.; SHEPHERD, R. J. Characterization of cassava vein mosaic virus: a distinct plant pararetrovirus. **Journal of General Virology,** v. 76, n. 5, 1271-1278, 1995.

CARVAJAL-YEPES, M.; OLAYA, C.; LOZANO, I.; CUERVO, M.; CASTAÑO, M.; CUELLAR, W. J. Unraveling complex viral infections in cassava (*Manihot esculenta* Crantz) from Colombia. **Virus Research**, v. 186, p. 76-86, 2014.

CHAPARRO-MARTÍNEZ, E. I.; TRUJILLO-PINTO, G. First report of cassava virus X in cassava in Venezuela. **Plant Disease,** v. 85, n. 10, p.1119, 2001.

CILIA, M.; PETER, K. A.; BEREMAN, M. S.; HOWE, K.; FISH, T.; SMITH, D.; GILDOW, F.; MACCOSS, M. J.; THANNHAUSER, T. W.; GRAY, S. M. Discovery and targeted LC-MS/MS of purified polerovirus reveals differences in the virus-host interactome associated with altered aphid transmission. **PLoS ONE,** v. 7, n. 10, p. e48177, 2012.

COLARICCIO, A.; E.B. RIVAS, E. B.; BONFIM JUNIOR, M. F.; PERUCH, L. A. M. Ocorrência do Cassava Common Mosaic Virus (CsCMV) e Cassava Vein Mosaic Virus (CsVMV) em cultivares de mandioca em Santa Catarina, Brasil. **Biológico**, v. 82, 1-9, 2020

COLLAVINO, A.; ZANINI, A. A.; MEDINA, R.; SCHALLER, S.; DI FEO, L. D. V. Cassava common mosaic virus infection affects growth and yield components of cassava plants (Manihot esculenta) in Argentina. **Plant Pathology**, v. 71, n. 4, p. 980-989, 2021.

COSTA, A. S. Observações sobre o mosaico comum e o mosaico das nervuras da mandioca (*Manihot utilissima* Pohl.). **Jornal de Agronomia**, v. 3, p. 239-248, 1940.

COSTA, T. M.; BLAWID, R.; ARANDA, M. A.; FREITAS, D.; ANDRADE, G. P.; INOUE-NAGATA, A. K.; NAGATA, T. Cucurbit aphid-borne yellows virus from melon plants in Brazil is an interspecific recombinant. **Archives of Virology,** v. 164, n. 1, p. 249-254, 2019.

CUELLAR, W. J.; DE SOUZA, J.; BARRANTES, I.; FUENTES, S.; KREUZE, J. F. Distinct cavemoviruses interact synergistically with sweet potato chlorotic stunt virus (genus

Crinivirus) in cultivated sweet potato. **The Journal of general virology**, v. 92, n. 5, p. 1233-1243, 2011.

CUENCA, M. A. G.; MANDARINO, D. C. **Aspectos agroeconômicos da cultura da mandioca**: características e evolução da cultura no Estado de Pernambuco entre 1990 e 2004. Aracaju: Embrapa Tabuleiros Costeiros, 2006. 24 p. (Documentos / Embrapa Tabuleiros Costeiros, 99).

DE SOUZA, J.; CUELLAR, W. 2011. Sequence analysis of the replicase gene of 'sweetpotato caulimo-like virus' suggests that this virus is a distinct member of the genus Cavemovirus. **Archives of Virology**, v. 156, p. 535-537, 2011.

DELFOSSE, V. C.; BARRIOS BARÓN, M. P.; DISTÉFANO, A. J. What we know about poleroviruses: Advances in understanding the functions of polerovirus proteins. **Plant Patholology,** v. 70, n. 5, p. 1047-1061, 2021.

DISTÉFANO, A. J.; KRESIC I. B.; HOPP, H. E. The complete genome sequence of a virus associated with cotton blue disease, cotton leafroll dwarf virus, confirms that it is a new member of the genus Polerovirus. **Archives of Virology**, v. 155, p.1849-1854, 2010.

EDGAR, R. C.; TAYLOR, J.; LIN, V.; ALTMAN, T.; BARBERA, P.; MELESHKO, D.; LOHR, D.; NOVAKOVSKY, G.; BUCHFINK, B.; AL-SHAYEB, B.; BANFIELD, J. F.; DE LA PEÑA, M.; KOROBEYNIKOV, A.; CHIKHI, R.; BABAIAN, A. Petabase-scale sequence alignment catalyses viral discovery. **Nature,** v. 602, n. 7895, p. 142-147, 2022.

ENI, A. O.; EFEKEMO, O. P.; ONILE-ERE, O. A.; PITA, J. S. South West and North Central Nigeria: Assessment of cassava mosaic disease and field status of African cassava mosaic virus and East African cassava mosaic virus. **Annals of Applied Biology**, v. 178, n. 3, p. 466-479, 2020.

ESPACH, Y.; MAREE, H. J.; BURGER, J. T. Complete genome of a novel endornavirus assembled from next-generation sequence data. **Journal of virology**, v. 86, n. 23, p. 13142, 2012.

FAO - Food and Agriculture Organization of the United Nations. **Production of Cassava: top 10 producers 2020.** Roma: Food and Agriculture Organization of the United Nations, 2016. Disponível em: < http://www.fao.org/faostat/en/#data/QC/visualize>. Acesso em: 10 jan. 2022.

FERRIOL, I.; SILVA-JUNIOR, D. M.; NIGG, J. C.; ZAMORA-MACORRA, E. J.; FALK, B. W. Identification of the cleavage sites of the RNA2-encoded polyproteins for two members of the genus Torradovirus by N-terminal sequencing of the virion capsid proteins. **Virology**, v. 496, p. 109-115, 2016.

GHOSH, S.; KANAKALA, S.; LEBEDEV, G.; KONTSEDALOV, S.; SILVERMAN, D.; ALON, T.; MOR, N.; SELA, N.; LURIA, N.; DOMBROVSKY, A.; MAWASSI, M.; HAVIV, S.; CZOSNEK, H.; GHANIM, M. Transmission of a new polerovirus infecting pepper by the whitefly *Bemisia tabaci.* **Journal of Virology**, v. 93, p. e00488-e00519, 2019.

GOODWIN, S.; McPHERSON, J. D.; McCOMBIE, W. R. Coming of age: ten years of next-generation sequencing Technologies. **Nature Reviews Genetics**, Londres, v. 17, p. 333-351, 2016.

GRAY, S.; GILDOW, F. E. Luteovirus-aphid interactions. **Annual Review of Phytopathology,** v. 41, p. 539-566, 2003.

GUO, J.; BOLDUC, B.; ZAYED, A. A.; VARSANI, A.; DOMINGUEZ-HUERTA, G.; DELMONT, T. O.; PRATAMA, A. A.; GAZITÚA, M. C.; VIK, D.; SULLIVAN, M. B.; ROUX, S. VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. **Microbiome**, v. 9, n. 1, p. 37, 2021.

HARRISON, B. D.; LENNON, A. M.; AITON, M. M. Properties of four previously uncharacterized cassava viruses. **Phytopathology**, v. 76, p. 1075, 1986.

HILLOCKS, R.; RAYA, M.; MTUNDA, K.; KIOZIA, H. Effects of brown streak virus disease on yield and quality of cassava in Tanzania**. Journal of Phytopathology**, v. 149, n. 7-8, p. 389-394, 2008.

HURWITZ, B. L.; HALLAM, S. J.; SULLIVAN, M. B. Metabolic reprogramming by viruses in the sunlit and dark ocean. **Genome Biology,** v. 14, n. R123, 2013.

IBGE - Instituto Brasileiro de Geografia e Estatística. **Levantamento Sistemático da Produção Agrícola:** Estatística da Produção Agrícola. Rio de Janeiro. Disponível em: < https://biblioteca.ibge.gov.br/visualizacao/periodicos/2415/epag_2021_jan.pdf> Acesso em: 03 fev. 2022.

JAAG, H. M.; KAWCHUK, L.; ROHDE, W.; FISCHER. R.; EMANS, N.; PRÜFER, D. An unusual internal ribosomal entry site of inverted symmetry directs expression of a potato leafroll polerovirus replication-associated protein. **Proceedings of the National Academy of Sciences of the United States of America**, v. 100, n. 15, p. 8939-8944, 2003.

JEEVALATHA, A.; KAUNDAL, P.; SHANDIL, R. K.; SHARMA, N. N.; CHAKRABARTI, S. K.; SINGH, B. P. Complete genome sequence of potato leafroll virus isolates infecting potato in the different geographical areas of India shows low level genetic diversity. **Indian Journal of vVirology,** v. 24, n. 2, p. 199-204, 2013.

JONES, P., DEVONSHIRE, J., DABEK, A., HOWELLS, C. First report of Cassava common mosaic potexvirus infecting Chaya (*Cnidoscolus chayamansa*) in Tuvalu. **Plant Disease,** v. 82, n. 5, p. 591, 1998.

KIDULILE, C. E.; MIINDA ATEKA, E.; ALAKONYA, A. E.; NDUNGURU, J. C. Efficacy of chemotherapy and thermotherapy in elimination of *East African cassava mosaic virus* from Tanzanian cassava landrace. **Journal of Phytopathology**, v. 166, n.10, p.739-745, 2018.

KING, A. M. Q.; ADAMS, M. J.; CARSTENS, E. B.; LEFKOWITZ, E. J. **Virus Taxonomy**: Ninth report of the International Committee on Taxonomy of Viruses. 9th. ed. San Diego: Elsevier Academic Press, 2012. 1272 p.

KITAJIMA, E. W.; COSTA, A. S. Rhabdovirus-like particles in tissues of five different plant species. **Fitopatologia Brasilera**, v. 4, n. 55, 1979.

KITAJIMA, E. W.; COSTA, A. S. Partículas esferoidais associadas ao virus do mosaico das nervuras da mandioca. **Bragantia**, v. 25, p. 211-221, 1966.

KITAJIMA, E. W.; WETTER, C. OLIVEIRA, A. R.; SILVA, D. M.; COSTA, A. S. Morfologia do vírus do mosaico comum da mandioca. **Bragantia**, v. 24, p. 247-260, 1965.

KREUZE, J. F.; VAIRA, A. M.; MENZEL, W.; CANDRESSE, T.; ZAVRIEV, S. K.; HAMMOND, J.; RYU, K. H.; ICTV, C. ICTV Virus Taxonomy Profile: Alphaflexivridae, **Journal of General Virology,** v. 101, p. 699-700, 2020.

LECOQ, H.; VERDIN, E.; TEPFER, M.; WIPF-SCHEIBEL, C.; MILLOT, P.; DAFALLA, G.; DESBIEZ, C. Characterization and occurrence of squash chlorotic leaf spot virus, a tentative new torradovirus infecting cucurbits in Sudan. **Archives of Virology,** v. 161, n. 6, p. 1651-1655, 2016.

LEFEUVRE, P.; MARTIN, D. P.; ELENA, S. F.; SHEPHERD, D. N.; ROUMAGNAC, P.; VARSANI, A. Evolution and ecology of plant viruses. **Nature Reviews Microbiology,** London, v.17, p. 632-644, 2019.

LEGG, J. P.; KUMAR, P. L.; MAKESHKUMAR, T.; TRIPATHI, L.; FERGUSON, M.; KANJU, E.; NTAWURUHUNGA, P.; CUELLAR, W. Cassava virus diseases: biology, epidemiology, and management. **Advances in Virus Research**, v. 91, n. 1, p. 86-131, 2014.

LENNON, A. M.; AITON, M. M.; HARRISON, B. D. Cassava viruses from South America. **Scottish Crop Research Institute Annual Report 1985**, n. 167, 1986.

LOZANO, I.; LEIVA, A. M.; JIMENEZ, J.; FERNANDEZ, E.; CARVAJAL-YEPES, M.; CUERVO, M.; CUELLAR, W. J. Resolution of cassava-infecting alphaflexiviruses: Molecular and biological characterization of a novel group of potexviruses lacking the TGB3 gene. **Virus Research**, v. 15, n. 241, p. 53-61, 2017.

MALIOGKA, V. I.; MINAFRA, A.; SALDARELLI, P.; RUIZ-GARCÍA, A. B.; GLASA, M.; KATIS, N.; OLMOS, A. Recent advances on detection and characterization of fruit tree viruses using high-throughput sequencing technologies. **Viruses**, v. 10, n. 8 p. 436-459, 2018.

MARTÍNEZ-GARCÍA, M.; SANTOS, F.; MORENO-PAZ, M.; PARRO, V.; ANTÓN, J. Unveiling viral-host interactions within the 'microbial dark matter'. **Nature communications**, v. 5, n. 4542, 2014.

MARTINS, M. L. L.; ORLANDINI, P.; MENDOZA F., J. M.; SILVEIRA, T. C. **Manihot in Flora do Brasil 2020**. Jardim Botânico do Rio de Janeiro. Disponível em: <http://reflora.jbrj.gov.br/reflora/floradobrasil/FB17600>. Acesso em: 27 jan. 2022.

MATTOS, P. L. P.; CARDOSO, E. M. R**. Cultivo da mandioca para o Estado do Pará.** Cruz das Almas: Embrapa Mandioca e Fruticultura, 2003. (Sistema de Produção, 13. Embrapa Mandioca e Fruticultura). Disponível em: http://sistemasdeproducao.cnptia.embrapa.br/FontesHTML/Mandioca/mandioca_para/index.htm. Acesso em: 24 maio 2022.

MAPA- Ministério da Agricultura, Pecuária e Abastecimento. **CultivarWeb**. Disponível em: <https://sistemas.agricultura.gov.br/snpc/cultivarweb/cultivares_registradas.php>. Acesso em: 03 mai. 2022.

MEJÍAS, A.; RODRÍGUEZ-ROMÁN, E.; ROMANO, M.; ZAMBRANO, K.; MARYS, E. New record of cassava common mosaic virus infecting chaya (*Cnidoscolus chayamansa* McVaug) in Venezuela. **Plant Disease**, v. 99, n. 8, p. 1190, 2015.

MCCALLUM, E. J.; ANJANAPPA, R. B.; GRUISSEM, W. Tackling agriculturally relevant diseases in the staple crop cassava (*Manihot esculenta*). **Current Opinion in Plant Biology**, v. 38, p. 50-58, 2017.

MIZUNO, C. M.; RODRIGUEZ-VALERA, F.; KIMES, N. E.; GHAI, R. Expanding the marine virosphere using metagenomics. **PLoS Genetics**, v. 9, n. 12, p. e1003987, 2013.

MOROZOV, S. Y.; SOLOVYEV, A. G. Triple gene block: Modular design of a multifunctional machine for plant virus movement. **Journal of General Virology**, v. 84, n. 6 p. 1351-1366, 2003. doi: 10.1099/vir.0.18922-0.

MULENGA, R. M.; LEGG, J. P.; NDUNGURU, J.; MIANO, D. W.; MUTITU, E. W.; CHIKOTI, P. C.; ALABI, O. J. Survey, molecular detection, and characterization of geminiviruses associated with Cassava Mosaic Disease in Zambia. **Plant Disease**, v. 100, n. 7, p. 1379-1387, 2016.

NATURE. **Nex-generation-sequencing**. London: Springer Nature, 2019. Disponível em: <https://www.nature.com/subjects/next-generation-sequencing>. Acesso em: 17 de out. de 2019.

NOLT, B.; PINEDA, B.; VELASCO, A. C. Surveys of cassava plantations in Colombia for virus and virus-like diseases. **Plant Patholology**, v. 41, p. 348-354, 1992.

OLIVEIRA, S. A. S.; FERREIRA, C. F.; DIAMANTINO, M. S. A. S.; SANTOS, T. A.; PEREIRA, J. D.; OLIVEIRA, E. J. First report of cassava torrado-like virus, cassava polero-like virus and cassava new alphaflexivirus associated with cassava frogskin disease in Brazil. **Journal of Plant Pathology**. v. 10, p. 247, 2020.

OLSEN, K. M. SNPs, SSRs and inferences on cassava's origin. **Plant Molecular Biology**, v. 56, n. 4, p. 517-526, 2004.

OWOR, B.; LEGG, J. P.; OKAO-OKUJA, G.; OBONYO, R.; OGENGA-LATIGO, M. W. The effect of cassava mosaic geminiviruses on symptom severity, growth and root yield of a cassava mosaic virus disease-susceptible cultivar in Uganda. **Annals of Applied Biology**, v.145, n. 3, p. 331-337, 2004.

PARK, M. R.; JEONG, R. D.; KIM, K. H. Understanding the intracellular trafficking and intercellular transport of potexviruses in their host plants. **Frontiers in plant science**. v. 5, n. 60, 2014. doi:10.3389/fpls.2014.00060

PECMAN, A.; KUTNJAK, D.; GUTIÉRREZ-AGUIRRE, I.; ADAMS, I.; FOX, A.; BOONHAM, N.; RAVNIKAR, M. Next generation sequencing for detection and discovery of plant viruses and viroids: comparison of two approaches. **Frontiers in Microbiology**, v. 8, p. 1998, 2017.

PINEDA, B.; JAYASINGHE, U.; LOZANO, J. C. La enfermedad "Cuero de Sapo" en yuca (*Manihot esculenta*, Crantz). **ASIAVA**, v 4, p. 10-12, 1983.

PLYUSNIN, I.; KANT, R.; JÄÄSKELÄINEN, A. J.; SIRONEN, T.; HOLM, L; VAPALAHTI, O.; SMURA, T. Novel NGS pipeline for virus discovery from a wide spectrum of hosts and sample types, **Virus Evolution**, v. 6, n. 2, p. veaa091, 2020.

REIS, N. A.; FONSECA, M. E. D. N.; RIBEIRO, S. G.; NAITO, F. Y. B.; BOITEUX, L. S.; PEREIRA-CARVALHO, R. D. C. Metagenomics of Neotropical Single-Stranded DNA Vvruses in tomato cultivars with and without the Ty-1 Gene. **Viruses**, v. 12, n.8, p. 819, 2020.

ROHWER, F.; THURBER, R. Viruses manipulate the marine environment. **Nature**, v. 459, p. 207-212, 2009.

ROTT, M. E.; KESANAKURTI, P.; BERWARTH, C.; RAST, H.; BOYES, I.; PHELAN, J.; JELKMANN, W. Discovery of negative-sense RNA viruses in trees infected with apple rubbery wood disease by Next-Generation Sequencing. **Plant Disease**, v. 102, n. 7, 1254-1263, 2018.

ROUX, S.; ENAULT, F.; HURWITZ, B. L.; SULLIVAN, M. B. VirSorter: Mining viral signal from microbial genomic data. **PeerJ**, v. 3, n. e985, 2015.

ROUX, S.; HALLAM, S. J.; WOYKE, T.; SULLIVAN, M. B. Viral dark matter and virus-host interactions resolved from publicly available microbial genomes. **eLife,** v. 4, n. e08490, 2015.

ROZADO-AGUIRRE, Z.; ADAMS, I.; COLLINS, L.; FOX, A.; DICKINSON, M.; BOONHAM, N. Detection and transmission of Carrot torrado virus, a novel putative member of the *Torradovirus* genus. **Journal of Virological Methods**, v. 235, p. 119-124, 2016.

SANTOS, J. A.C. M.; SILVA, A. M. F. S.; DE MESQUITA, J. C. P.; BLAWID, R. Transcriptomic analyses reveal highly conserved plant amalgavirus genomes in different species of *Allium*. **Acta Virologica,** v. 66, n. 1, p. 11-17, 2022.

SEO, J. K.; KANG, M.; KWAK, H. R.; KIM, M. K.; KIM, C. S.; LEE, S. H.; KIM, J. S.; CHOI, H. S. Complete genome sequence of motherwort yellow mottle virus, a novel putative member of the genus *Torradovirus*. **Archives of Virology**, v. 160, n. 2, p. 587-590, 2015.

SIMMONDS, P.; ADAMS, M. J.; BENKŐ, M.; BREITBART, M.; BRISTER, J. R.; CARSTENS, E. B.; DAVISON, A. J.; DELWART, E.; GORBALENYA, A. E.; HARRACH, B.; HULL, R.; KING, A. M.; KOONIN, E. V.; KRUPOVIC, M.; KUHN, J. H.; LEFKOWITZ, E. J.; NIBERT, M. L.; ORTON, R.; ROOSSINCK, M. J.; SABANADZOVIC, S.; SULLIVAN, M. B.; SUTTLE, C. A.; TESH, R. B.; VAN DER VLUGT, R. A.; VARSANI, A.; ZERBINI, F. M. Consensus statement: Virus taxonomy in the age of metagenomics. **Nature Reviews Microbiology,** v. 15, n. 3, p. 161-168, 2017.

SÕMERA, M.; FARGETTE, D.; HÉBRARD, E.; SARMIENTO, C. (2021). ICTV Virus Taxonomy Profile: Solemoviridae 2021. **The Journal of General Virology**, v. 102, n. 12, p. 001707, 2021.

SMIRNOVA, E.; FIRTH, A. E.; MILLER, W. A.; SCHEIDECKER, D.; BRAULT, V.; REINBOLD, C.; RAKOTONDRAFARA, A. M.; CHUNG, B. Y. W.; ZIEGLER-GRAFF, V. Discovery of a Small Non-AUG-Initiated ORF in Poleroviruses and Luteoviruses That Is Required for Long-Distance Movement. **PLoS Pathogens**, v. 11, n. 5, p. e1004868, 2015.

SONENBERG, N.; SHATKIN, A. J.; RICCIARDI, R. P.; RUBIN, M.; GOODMAN, R. M. Analysis of terminal structures of RNA from potato virus X. **Nucleic Acids Research**, v. 5, n. 7, p. 2501-2512, 1978.

TEYCHENEY, P.; ANDREW D.W. GEERING, I. D.; HULL, R.; KREUZE, J. F.; MULLER, B. L. E.; OLSZEWSKI, N.; PAPPU, H.; POOGGIN, M. M.; ROCHERT-POGGELER, K. R.; SCHOELZ, J. E.; SEAL, S.; STAVOLONE, L.; UMBER, M.; ICTV, R. C. ICTV Virus Taxonomy Profile: *Caulimoviridae,* **Journal of General Virology, , v.** 101, p. 1025-1026, 2020

THOMPSON, J.R., DASGUPTA, I., FUCHS, M., IWANAMI, T., KARASEV, A.V., PETRZIK, K., SANFAC¸ON, H., TZANETAKIS, I. E., VAN DER VLUGT, R., WETZEL, T., YOSHIKAWA, N.; ICTV, R. C. ICTV virus taxonomy profile: *Secoviridae*. **Journal of General Virology**, v. 98, p. 529-531, 2017.

TOMLINSON, K. R.; BAILEY, A. M.; ALICAI, T.; SEAL, S.; FOSTER, G. D. Cassava brown streak disease: historical timeline, current knowledge and future prospects. **Molecular Plant Pathology**, v.19, n. 5, p. 1282-1294, 2018.

TUO, D. C.; ZHAO, G. Y.; YAN, P.; LI, R. M.; CHEN, X.; WANG, W. Q.; LI, X. Y.; YAO, Y.; CUI, H. G.; SHEN, W. T.; ZHOU, P. First report of cassava common mosaic virus infecting cassava in Mainland China. **Disease Notes**, v. 104, n. 3, p. 1-9, 2020.

VAN DER VLUGT, R. A.; VERBEEK, M.; DULLEMANS, A. M.; WINTERMANTEL, W. M.; CUELLAR, W. J.; FOX, A. THOMPSON, J. R. Torradoviruses. **Annual Review of Phytopathology**, v. 53, p. 485-512, 2015.

VENTURINI, M. T.; ARAÚJO, T. S.; ABREU, E. F. M.; ANDRADE, E. C.; SANTOS, V. S.; SILVA, M. R.; OLIVEIRA, E. J. Crop losses in Brazilian cassava varieties induced by the Cassava common mosaic virus. **Scientia Agricola**, v. 73, n. 6, p. 520-524, 2016.

VERBEEK M.; DULLEMANS, A. M.; VAN DEN HEUVEL, J. F.; MARIS, P. C.; VAN DER VLUGT.; R. A. Identification and characterisation of tomato torrado virus, a new plant picorna-like virus from tomato. **Archives of Virology,** v. 152, n. 5, p. 881-890, 2007.

VERBEEK, M.; DULLEMANS, A.; VAN DEN HEUVEL, J.; MARIS, P. C.; VAN DER VLUGT, R. A. A. Tomato marchitez virus, a new plant pi-corna-like virus from tomato related to tomato torrado virus. **Archives of Virology,** v. 153, n. 1, p. 127-134, 2008.

VERBEEK, M.; DULLEMANS, A. M.; VAN RAAIJ H. M.; VERHOEVEN, J. T.; VAN DER VLUGT, R. A. Lettuce necrotic leaf curl virus, a new plant virus infecting lettuce and a proposed member of the genus *Torradovirus*. **Archives of Virology,** v. 159, n. 4, p. 801-805, 2014.

VERBEEK, M.; DULLEMANS, A.; VAN DEN HEUVEL, H.; MARIS, P.; VAN DER VLUGT, R. Tomato chocolàte virus: A new plant virus infecting tomato and a proposed member of the genus Torradovirus. **Archives of Virology,** v. 155, n. 5, p. 751-755, 2010.

VERBEEK, M.; VAN BEKKUM, P. J.; DULLEMANS, A. M.; VAN DER VLUGT, R. A. A. Torradoviruses are transmitted in a semi-persistent and stylet-borne manner by three whitefly vectors. **Virus Research**, v. 186, p. 55-60, 2014.

VILLAMOR, D.; HO, T.; AL RWAHNIH, M.; MARTIN, R.; TZANETAKIS, I. High throughput sequencing for plant virus detection and discovery. **Phytopathology**, v. 109, n. 5, p. 716-725, 2019.

VILLAMOR, D. E. V.; MEKURIA, T. A.; PILLAI, S. S.; EASTWELL, K. C. High-Throughput sequencing identifies novel viruses in nectarine: Insights to the etiology of stem-pitting disease. **Phytopathology,** v. 106, p. 519-527, 2016.

YU, N.; YANG, Y.; LI, J.; LI, W.; LIU, Z. Complete genome sequence of a distinct isolate of cassava common mosaic virus (CsCMV) infecting cassava in Hainan, China. **European Journal of Plant Pathology,** v. 158, p. 583-587, 2020.

WALTER, B.; LADEVEZE, I.; ETIENNE, L.; FUCHS, M. Some properties of a previously undescribed virus from cassava: Cassava American Latent Virus**. Annals of Applied Biology**, v. 115, p. 279-289, 1989.

WINTER, S.; KOERBLER, M.; STEIN, B.; PIETRUSZKA, A.; PAAPE, M.; BUTGEREITT, A. Analysis of cassava brown streak viruses reveals the presence of distinct

virus species causing cassava brown streak disease in East Africa. **Journal of General Virology**, v. 91, n. 5, p. 1365-1372, 2010.

WINTERMANTEL, W. M.; HLADKY, L. L.; CORTEZ, A. A. Genome sequence, host range, and whitefly transmission of the torrado-virus Tomato necrotic dwarf virus. **Acta Horticulturae,** v. 1207, p. 295-302, 2018.

ZHAO, K.; YIN, Y.; HUA, M.; WANG, S.; MO, X.; YUAN, E.; ZHENG, H.; LIN, L.; CHEN, J.; LU, Y.; CHEN, H.; LU, Y.; CHEN, J.; PENG, J.; YAN, F. Pod pepper vein yellows virus, a new recombinant polerovirus infecting *Capsicum frutescens* in Yunnan province, China. **Virology Journal**, v. 18, n. 42, 2021.

ZANINI, A. A.; CUELLAR, W. J.; CELLI, M. G.; LUQUE, A. V.; MEDINA, R. D.; CONCI, V. C.; DEL DI FEO, L. V. Distinct strains of the re-emergent cassava common mosaic virus (genus: *Potexvirus*) infecting cassava in Argentina. **Plant Pathology,** v. 67, n. 8, p. 1814-1820, 2018.

ZHENG, L.; CAO, M.; WU, L.; LIU, H.; CHEN, M.; LI, R. First identification and molecular characterization of a novel cavemovirus infecting *Epiphyllum* spp. **Archives of Virology**, v. 165, n. 9, p. 2083-2086, 2020.

ZINGA, I.; CHIROLEU, F.; LEGG, J.; LEFEUVRE, P.; KOMBA, E. K.; SEMBALLA, S.; YANDIA, S. P.; MANDAKOMBO, N. B.; REYNAUD, B.; LETT, J. Epidemiological assessment of cassava mosaic disease in Central African Republic reveals the importance of mixed viral infection and poor health of plant cuttings. **Crop Protection**, v. 44, p. 6-12, 2013.

CAPÍTULO II

Sequence diversity analysis of torrado-like viruses in cassava from Brazil reveals a novel member of the genus *Torradovirus*

Plants

Sequence diversity analysis of torrado-like viruses in cassava from Brazil reveals a novel member of the genus *Torradovirus*

Géssyka Rodrigues de Albuquerque[1], Alessandra de Jesus Boari[2], Eder Jorge de Oliveira[3], Paolo Margaria[4], Stephan Winter[4] & Rosana Blawid[1*]

[1] Universidade Federal Rural de Pernambuco, Recife, Pernambuco, Brazil.

[2] Embrapa Amazônia Oriental, Belém, Pará, Brazil.

[3] Embrapa Mandioca e Fruticultura, Cruz das Almas, Bahia, Brazil.

[4] Leibniz Institut DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Braunschweig, Germany

*Correspondence: rblawid@gmail.com

Abstract: As indicated by previous studies, the overall viral diversity in cassava (*Manihot esculenta*) is largely underestimated especially in Neotropical areas. Here, we sequenced 29 Illumina libraries aiming to provide insights into the incidence and sequence diversity of cassava-infecting members of the emergent viral genus *Torradovirus* (family *Secoviridae*) occurring in Brazil. Eleven nearly-complete genome sequences of cassava torrado-like viruses (CsTLVs) were assembled, and sequences from the protease-polymerase (Pro-Pol) and the combined coat proteins (CP-A, CP-B, CP-C) regions were used in Bayesian phylogenetic analyses. Two species of torrado-like virus were found, one grouping with the formerly reported CsTLV from Colombia, and one separate clade composed of a putative new species, for which the name cassava torrado-like virus 2 (CsTLV-2) is proposed. CsTLV-2 shared less than 66.3% amino acid identity in the Pro-Pol region and less than 66.4% identity in the combined CPs region compared to the known species. Analysis of a RNAseq dataset of a cassava sample from Colombia available in GenBank, revealed the presence of a CsTLV more closely related to the Brazilian virus than to the formerly reported Colombian sequences.

Keywords: novel cassava torrado-like virus; *Manihot esculenta*; Illumina; high-throughput sequencing

1. Introduction

Cassava (*Manihot esculenta* Crantz) is an important traditional food crop in tropical regions across Africa, Asia, and South America. The impact of virus-caused cassava diseases is a serious concern since they can harm millions of people who depend on this crop for subsistence agriculture (McCallum *et al*., 2017). Unlike Africa, where geminivirids and potyvirids are the major pathogens affecting cassava cultivation, such viruses have not been reported in this crop in Latin America. However, an array of RNA and DNA viruses belonging to the families *Alphaflexiviridae*, *Secoviridae*, *Luteoviridae, Reoviridae* among RNA viruses and to the *Caulimoviridae* among DNA viruses, have been described infecting cassava in South America (Oliveira *et al*., 2020, Carvajal-Yepes, *et al*., 2014; Calvert *et al*., 2008). The Cassava Vein Mosaic Disease (CVMD) and Cassava Common Mosaic Disease (CCMD) were the first diseases to be reported in cassava in South America (Costa, 1940, Silva *et al*., 1963), although CCMD is not considered a major issue. The most economically important disease in Latin America is the Cassava Frog Skin Disease (CFSD) syndrome. While the disease has been known for many years, with a first serious outbreak reported in 1971 in the Cauca region of Colombia (Hernández *et al*., 1975; Pineda et al., 1983), the precise etiology has yet to be determined. The disease is of serious concern for cassava cultivation in different countries, such as Brazil, Costa Rica, Panama, Peru and Venezuela. Viruses as well as phytoplasmas (group 16SrIII) have been reported from plant material exhibiting CFSD-like symptoms (Calvert et al., 2008; Alvarez et al., 2009; Carvajal-Yepes et al., 2014).

The main viruses reported in cassava in Brazil are cassava common mosaic virus (CsCMV), cassava vein mosaic virus (CsVMV) (Costa, 1940), cassava symptomless virus (CsSLV) (Kitajima & Costa, 1979), cassava American latent virus (CsALV) (Walter *et al*., 1989), cassava polero-like virus (CsPLV), cassava new alphaflexivirus (CsNAV) and cassava torrado-like virus (CsTLV) (Oliveira *et al.,* 2020). Mixed infections often occur (Oliveira *et al.,* 2020). The CsTLV was first identified in cassava in Colombia (Carvajal-Yepes *et al*., 2014). The virus is a proposed member of the genus *Torradovirus*, family *Secoviridae*. For reported and accepted species by ICTV, the virions consist of non-enveloped isometric particles and contain a genome composed of two single-stranded positive-strand RNAs, RNA1 and RNA2 in the range of 6911-7802 nucleotides (nt) and 5695-4701 nt in length, respectively for reported and accepted species by ICTV. The RNA1 codes for a polyprotein of 2151-2224 amino acids (aa), which is processed into a protease cofactor, a helicase, a 3C-like protease and an RNA-dependent RNA polymerase (RdRp). The RNA2 encodes a hypothetical protein (185-212 aa) and a polyprotein (1168-1223 aa), that is processed into a movement protein (MP) and three coat proteins (CP-A, CP-B, and CP-C). The genomic RNAs of the torradoviruses usually have a long 3'-UTR and a poly(A)-tail (Thompson *et*

al., 2017; Ferriol *et al*., 2016; van der Vlugt *et al.,* 2015). Transmission vectors have not yet been identified (van der Vlugt *et al.,* 2015).

Although cassava torrado-like virus sequences have been reported from Brazil (Oliveira *et al*., 2020), complete viral genomes and sequence diversity have not been so far determined. The high degree of sequence variability previously observed among cassava torrado-like viruses and the inability to detect isolates collected in the same region (Carvajal-Yepes *et al*., 2014), showed that a high degree of genetic diversity exists. In this work, we investigated the occurrence and diversity of torrado-like viruses in cassava material from Brazil. Our investigation revealed a novel torrado-like virus species, for which the name cassava torrado-like virus 2 (CsTLV-2) is proposed.

2. Results

2.1. Analysis of HTS datasets

An overview of the datasets generated in this work is presented in Table1. The number of total reads per sample ranged from 362,290 to 48,781,540 for MiSeq sequencing and from 6,656,458 to 98,271,182 for libraries processed with NextSeq-500. The percentage of trimmed reads ranged between 0,6 and 18%; the number of contigs per sample generated by SPAdes ranged between 4,135 and 253,292. The HiSeq 2500 sequencing generated a total number 17,254,734 reads and 15,147,828 after trimming, of which 56,524 were assembled into contigs using SPAdes. In 56.7 % of the HTS datasets (from 30), tBLASTX alignments revealed the presence of sequences corresponding to CsTLV (Table S1). From those, eleven torrado-like virus nearly-complete genomes could be assembled and were considered in further analyses. Analysis of the dataset retrieved from GenBank also resulted in assembly of four contigs having CsTLV as closest relative.

2.2. Characterization and genome structure of the CsTLVs

Alignment analyses using BLASTp revealed that the sequences assembled from five datasets were close to the CsTLV previously reported from Colombia (Carvajal-Yepes et al., 2014), and are hereafter named CsTLV-1. The assembled genomic sequences of this CsTLV-1 were in the range 7085-7251 nt for RNA1 and 4366-4469 nt for RNA2, excluding the poly(A) tail. The ORF1 of RNA1 was 7011 nt in length and encoded a predicted polyprotein of 2337 aa. The ORF1 of RNA2 was 681 nt in length, coding for a putative protein of 227 aa, while the ORF2 consisted of 3540 nt, coding for a polyprotein of 1180 aa (Table 1). The complete genomic sequence of isolate PV-1279, from sample P26_2E, chosen as reference collection isolate for this species. The sequence

showed 98.7% aa acid identity in the Pro-Pol region and 98.9% in the combined CPs region to the already described CsTLV from Colombia (GenBank acc. MF449522, MF449523).

Table 1. Length of the assembled cassava torrado-like virus sequences, untranslated regions and predicted open reading frames (ORFs).

	Virus	CsTLV-1*	CsTLV-1	CsTLV-1	CsTLV-1	CsTLV-1	CsTLV-1	CsTLV-2	CsTLV-2	CsTLV-2	CsTLV-2	CsTLV-2
	Dataset	P26_2E	DSC407	DSC492	DSCG4	DSCG5	SRR1050897	AJB1_S86	P26_P27_4C	P27_6F	P27_6G_1	P27_6G_2
RNA1	Whole contig -poly (A)	7251	7142	7154	7135	7085	7129	7402	7451	7421	7451	7284
	5'UTR	148	62	62	62	62	62	155	173	143	173	173
	ORF1	7011	7011	7011	7011	7011	7011	6891	6891	6891	6891	6891
	3'UTR	92	69	81	62	12	56	356	387	387	387	220
RNA2	Whole contig -poly (A)	4469	4392	4385	4383	4394	4366	4253	4593	4583	4582	4593
	5'UTR	140	75	75	75	75	71	119	90	90	90	90
	ORF1	681	681	681	681	681	681	660	660	660	660	660
	ORF2	3540	3540	3540	3540	3540	3540	3432	3432	3432	3432	3432
	3'UTR	154	143	136	134	145	121	89	458	448	447	458

*verified by RACE

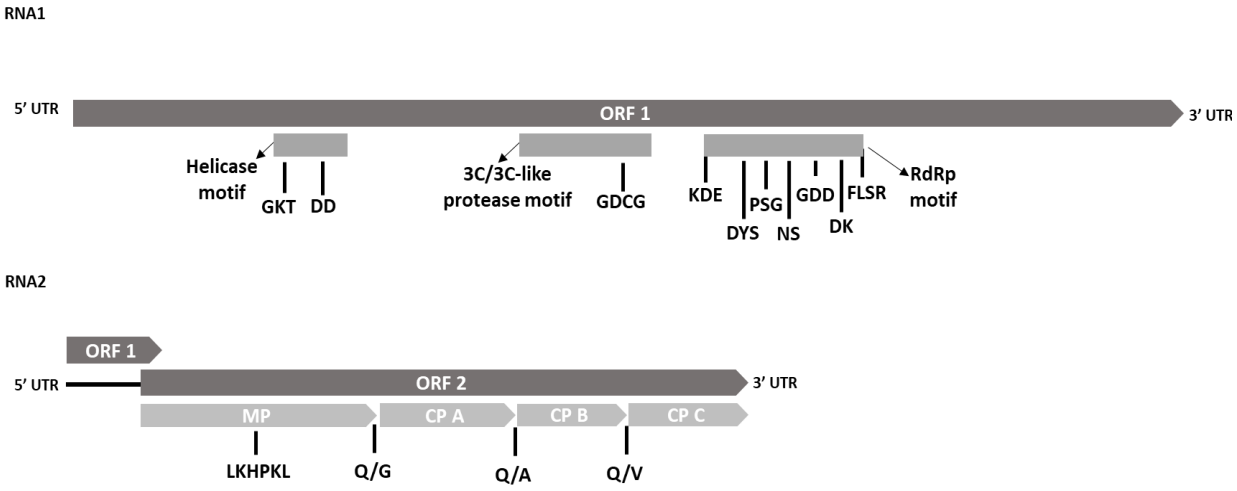
Alignment of other torrado-like virus sequences, assembled from datasets AJB1_S86, P26_ P27_4C, P27_6F and P27_6G, revealed aa identity below 66.1% in the Pro-Pol region and 66.9% in the combined CPs region (Table 2) compared to the previously determined CsTLV. According to the species demarcation criteria for members of the *Secoviridae* family, the obtained genomic sequences should be assigned to a new species, for which we propose the name "cassava torrado-like virus 2" (CsTLV-2). The assembled RNA1 segment of CsTLV-2 was 7284-7451 nt long, excluding the poly(A) tail, while the RNA2 ranged from 4553 to 4593 nt. The RNA1 ORF1 was 6891 nt, coding for a polyprotein of 2297 aa. The ORF1 of RNA2 was 660 nt in length, encoding a putative protein of 220 aa, and the ORF2 of RNA2 consisted of 3432 nt, coding for a polyprotein of 1144 aa.

Analysis of characteristics aa motifs of CsTLVs, showed the presence of conserved motifs in the polyprotein encoded by RNA1: the GKT and DD motifs in the helicase domain and KDE, DYS, NS, PSG, GDD, DK, FLSR motifs in the predicted RNA-dependent RNA-polymerase (Gaafar & Ziebell, 2019; Bruenn, 2003; Poch et al., 1989) were common to all CsTLV sequences (Figure 1a). Instead, the 3C-like protease motifs consisted of GDCG/GECG in CsTLV-1 and 2 respectively, with an isoleucine or glutamine residue at the substrate binding site (Figure 1b). In the RNA2-encoded polyprotein sequence the expected LKHPKL/LKHPRL movement protein motif (van der Vlugt et al., 2015) was present.

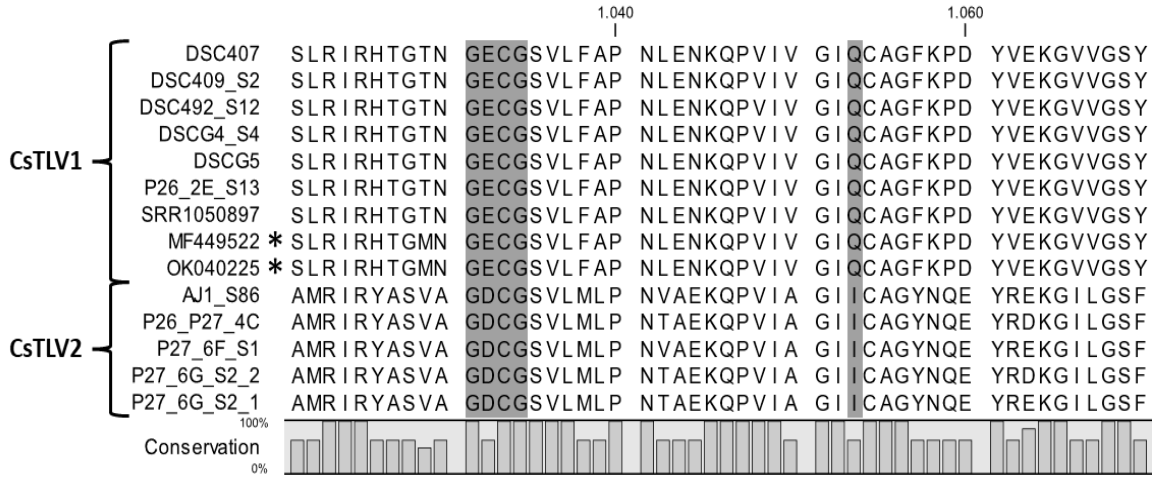
The predicted molecular weight (MW) of the hypothetical protein encoded by ORF1 of RNA2 was ~ 25.2 kDa for CsTLV-1 and ~24.4 kDa for CsTLV-2. The encoded MP was predicted to be ~55.3 kDa for CsTLV-1 and ~52.5 kDa for CsTLV-2. Based on multiple sequence alignment with carrot torradovirus, the predicted polyprotein cleavage sites were determined as Q_{484}/G_{485}, Q_{709}/A_{710} and Q_{963}/V_{964}, and Q_{462}/G_{463}, Q_{686}/A_{687} and $Q_{911}/V_{912,}$ in CsTLV-1 and CsTLV-2 respectively. The putative MW of CP-A, CP-B and CP-C were predicted as ~24.2, 27.7 and 23.9 kDa for CsTLV-1, and 24.3, 24.6, and 26.0 kDa for CsTLV-2.

Figure 1. Genomic structure of cassava torrado-like virus 2 (CsTLV-2) and alignment of the 3C-like protease region. (a) Schematic representation of the genome organization of CsTLV-2 (from dataset P26_P27_4C). Dark gray boxes show positions of open reading frames (ORFs) and protein motifs. The sequence of the 5´ and 3´ regions have not been verified by RACE. (b) Multiple alignments of the 3C-like protease motifs in the polyprotein encoded by RNA1, using QIAGEN CLC Genomics.

(a)



(b)



* GenBank accession

2.3. Sequence diversity

Sequence comparisons of CsTLV-1 from Brazil showed percentage of nucleotide identities with other CsTLV-1 sequences, ranging from 96.0 to 99.9% in the Pro-Pol region. The percentage of nt and aa identity of CsTLV-1 (P26_2E) with the sequence from Colombia (SRR1050897, assembled in this work) was of 96.8% and 100%, respectively, and of 87.8% and 98.7% with the reference sequence from Colombia (MF449522) in the Pro-Pol region. The CsTLV-2 sequences shared 81.7% to 99.7% identities at nt level and 95.6-100% at aa level in the Pro-Pol region among them. The Pro-Pol region of the Brazilian CsTLV2 from Amazon (AJB1_S86) showed 82.3-85% nt and 95.6-97.8-99% aa identities with all other sequences from Bahia, Brazil.

Comparisons of the CPs region showed that CsTLV-1 sequences from Brazil shared 97-100% nt identity with each other, 97.5-98.4% with the assembled sequence from Colombia (SRR1050897) and 89.7-90.0% with the reference sequence from GenBank (MF449523). At aa level, the CPs region showed identities of 99.7-100% among the Brazilian sequences and of 98.7-99.0% with the assembled Colombian sequence (SRR1050897). Among the CP CsTLV-2 sequences the % of nt and aa identity was of 87-99.9 and 98.7-100.0, respectively. The CPs region of CsTLV-2 from Amazon (AJB1_S86) showed 87-87.8% nt and 98.7-99% aa identities with other Brazilian sequences. Interspecies comparisons of CsTLV-1 and CsTLV-2 sequences revealed identities varying from 37.9 to 58.6% of nt and 53.1-53.9% of aa of the complete RNA1 polyprotein region (Table S2). In the Pro-Pol region, nt identities of 62.9-65.6 and of 65.6-66.3% aa identities were observed (Table 2). Comparisons at nucleotide and concatenated proteins sequences of the RNA2 region showed that nt CsTLV-1 sequence identities varied from 91.3% to 99.9%, while CsTLV-2 varied from 98.2% to 100% at the aa level (Table S3).

Table 2. Distance matrix of MAFFT based comparison (SDT analysis) of the Pro-Pol region from RNA1 and combined CPs region from RNA2.

	Pro-pol nt	DSC407	DSC409	DSC492	DSCG4	DSCG5	P26_2E	SRR1050897	MF449522	OK040226	AJB1_S86	P26_P27_4C	P27_6F	P27_6G_1	P27_6G_2
CsTLV-1	DSC407	100,0													
	DSC409	96,8	100,0												
	DSC492	96,8	96,0	100,0											
	DSCG4	98,3	96,3	96,3	100,0										
	DSCG5	98,1	96,6	96,3	98,5	100,0									
	P26_2E	96,6	96,0	99,9	96,2	96,1	100,0								
	SRR1050897	97,4	97,8	97,0	97,3	97,3	96,8	100,0							
	MF449522	88,0	88,0	88,0	87,8	87,5	87,8	88,0	100,0						
	OK040225	88,0	87,8	87,8	87,8	87,7	87,7	87,9	99,4	100,0					
CsTLV-2	AJB1_S86	63,6	63,9	63,9	63,9	63,7	63,9	64,3	64,0	63,9	100,0				
	P26_P27_4C	63,6	63,4	63,8	63,1	62,9	63,8	63,6	63,9	63,6	82,3	100,0			
	P27_6F	65,1	64,5	64,9	65,1	65,0	64,8	65,6	65,0	64,8	85,0	81,7	100,0		
	P27_6G_1	63,9	63,9	64,2	63,5	63,5	64,2	64,5	65,0	64,8	84,0	92,4	89,0	100,0	
	P27_6G_2	63,6	63,4	63,8	63,1	62,9	63,8	63,6	64,0	63,8	82,3	99,7	81,8	92,7	100,0
	Pro-pol aa	DSC407	DSC409	DSC492	DSCG4	DSCG5	P26_2E	SRR1050897	MF449522	OK040226	AJB1_S86	P26_P27_4C	P27_6F	P27_6G_1	P27_6G_2
CsTLV-1	DSC407	100,0													
	DSC409	100,0	100,0												
	DSC492	100,0	100,0	100,0											
	DSCG4	100,0	100,0	100,0	100,0										
	DSCG5	100,0	100,0	100,0	100,0	100,0									
	P26_2E	100,0	100,0	100,0	100,0	100,0	100,0								
	SRR1050897	100,0	100,0	100,0	100,0	100,0	100,0	100,0							
	MF449522	98,7	98,7	98,7	98,7	98,7	98,7	98,7	100,0						
	OK040225	99,1	99,1	99,1	99,1	99,1	99,1	99,1	99,6	100,0					
CsTLV-2	AJB1_S86	66,1	66,1	66,1	66,1	66,1	66,1	66,1	66,3	66,1	100,0				
	P26_P27_4C	65,6	65,6	65,6	65,6	65,6	65,6	65,6	65,9	65,6	95,6	100,0			
	P27_6F	66,1	66,1	66,1	66,1	66,1	66,1	66,1	66,3	66,1	97,8	96,3	100,0		

	P27_6G_1	66,1	66,1	66,1	66,1	66,1	66,1	66,1	66,3	66,1	96,3	98,5	97,8	100,0	
	P27_6G_2	65,6	65,6	65,6	65,6	65,6	65,6	65,6	65,9	65,6	95,6	100,0	96,3	98,5	100,0
	CP nt	DSC407	DSC409	DSC492	DSCG4	DSCG5	P26_2E	SRR1050897	MF449522	OK040226	AJB1_S86	P26_P27_4C	P27_6F	P27_6G_1	P27_6G_2
CsTLV-1	DSC407	100,0													
	DSC409	98,9	100,0												
	DSC492	97,1	97,2	100,0											
	DSCG4	98,9	98,5	97,0	100,0										
	DSCG5	98,9	98,6	97,2	98,7	100,0									
	P26_2E	97,1	97,2	100,0	96,9	97,1	100,0								
	SRR1050897	98,4	98,1	97,5	98,0	98,2	97,5	100,0							
	MF449523	89,9	89,7	89,8	89,9	90,0	89,8	89,9	100,0						
	OK040226	97,9	97,7	97,2	97,7	97,9	97,2	99,0	89,7	100,0					
CsTLV-2	AJB1_S86	63,8	63,6	64,1	64,2	64,2	64,1	63,9	63,6	64,0	100,0				
	P26_P27_4C	63,5	63,8	63,4	63,7	63,6	63,4	63,6	64,9	63,7	87,7	100,0			
	P27_6F	63,5	63,9	63,5	63,7	63,7	63,5	63,7	64,9	63,8	87,8	99,9	100,0		
	P27_6G_1	63,9	63,9	64,0	63,9	64,1	64,0	64,1	64,7	64,1	87,0	94,0	94,0	100,0	
	P27_6G_2	63,4	63,8	63,4	63,6	63,7	63,4	63,6	64,9	63,8	87,7	99,8	99,6	93,9	100,0
	CP aa	DSC407	DSC409	DSC492	DSCG4	DSCG5	P26_2E	SRR1050897	MF449522	OK040226	AJB1_S86	P26_P27_4C	P27_6F	P27_6G_1	P27_6G_2
CsTLV-1	DSC407	100,0													
	DSC409	99,9	100,0												
	DSC492	99,7	99,7	100,0											
	DSCG4	99,9	99,9	99,9	100,0										
	DSCG5	99,9	99,9	99,9	100,0	100,0									
	P26_2E	99,7	99,7	100,0	99,9	99,9	100,0								
	SRR1050897	99,7	99,7	100,0	99,9	99,9	100,0	100,0							
	MF449523	98,7	98,7	99,0	98,8	98,8	99,0	99,0	100,0						
	OK040226	99,7	99,7	100,0	99,9	99,9	100,0	100,0	99,0	100,0					
CsTLV-2	AJB1_S86	66,6	66,6	66,7	66,6	66,6	66,7	66,7	66,3	66,7	100,0				
	P26_P27_4C	66,7	66,7	66,9	66,7	66,7	66,9	66,9	66,4	66,9	99,0	100,0			
	P27_6F	66,7	66,7	66,9	66,7	66,7	66,9	66,9	66,4	66,9	99,0	100,0	100,0		

P27_6G_1	66,7	66,7	66,9	66,7	66,7	66,9	66,9	66,4	66,9	98,7	99,7	99,7	100,0	
P27_6G_2	66,7	66,7	66,9	66,7	66,7	66,9	66,9	66,4	66,9	99,0	100,0	100,0	99,7	100,0

2.4. Phylogenetic analyses

Bayesian phylogenies were calculated on the Pro-Pol region and on the concatenated ORF1-ORF2 aa region of RNA2 from CsTLV-1 and CsTLV-2. The average standard deviation of split frequencies was of 0.002254 using the Pro-Pol region and of 0.002710 with the RNA2 region, indicating a good convergence. Phylogenetic tree reconstructions revealed two distinct clades. The reference sequence from Colombia (Carvajal-Yepes et al., 2014) grouped with the newly assembled CsTLV sequence from Colombia (SRR1050897) and six Brazilian sequences (DSC407, DCS409, DSC492, DSCG4, DSCG5 and P26_2E). The second clade grouped only CsTLV sequences from Brazil assigned to the proposed CsTLV-2 (Figure 2, Figure 3).

Figure 2. Bayesian phylogenetic reconstruction of torradoviruses based on the amino acid alignment of the Pro-Pol region (467 aa of aligned sites based on the CsTLV Sec13 sequence – GenBank acc.: MF449522). The scale bar represents 0.2 substitutions per amino acid position.

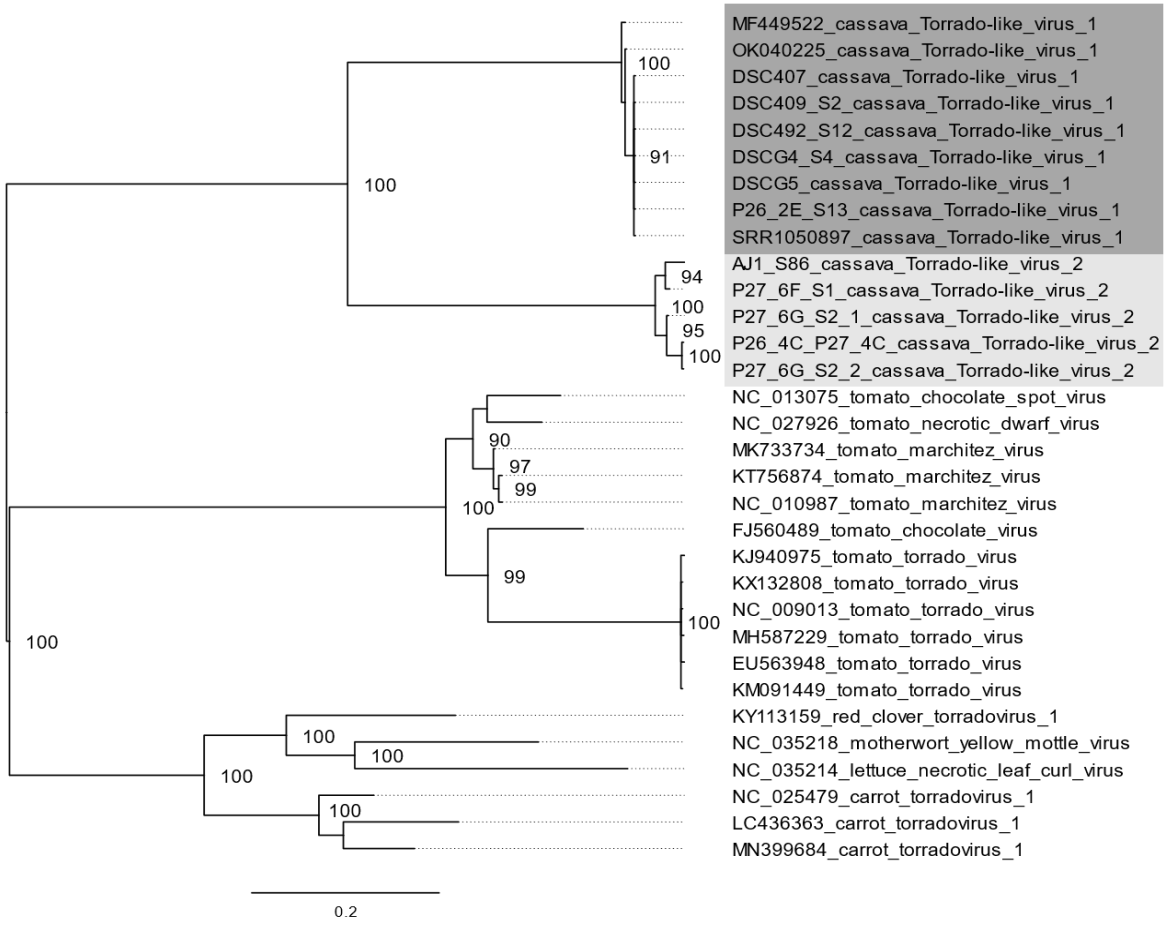
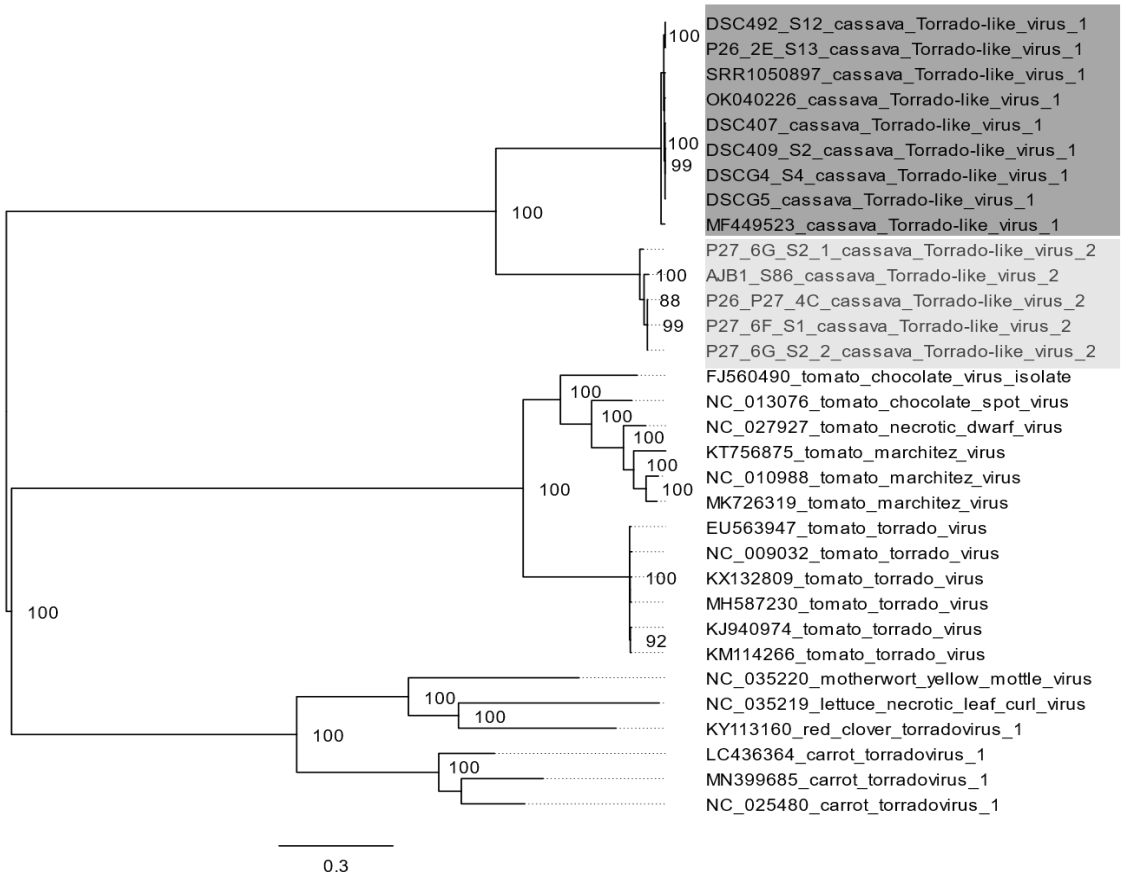


Figure 3. Bayesian phylogenetic reconstruction of torradoviruses based on the amino acid sequence alignment of the ORF1-ORF2 region (1546 aa of aligned sites based on the CsTLV Sec13 sequence - accession number: MF449523). The scale bar represents 0.4 substitutions per amino acid position.



2.5 Detection of CsTLV in cassava propagation material

In vitro propagation material from a collection of cassava lines, whose sanitary status was previously determined by HTS, was tested for the presence of CsTLVs by RT-PCR, using primers designed to specifically detect CsTLV-1 and CsTLV-2. The established RT-PCR assays could specifically detect CsTLV-1 and CsTLV-2 in the infected material, with no cross-reaction among the species, neither amplification of specific targets in the healthy controls.

3. Discussion

The history of torradoviruses (genus *Torradovirus*, family *Secoviridae*) began in 2007 with the description of the type member of the genus, *Tomato torrado virus* (Verbeek et al., 2007). Since then, new species have been discovered in different hosts, such as carrot, lettuce, motherwort,

squash (Wintermantel et al., 2018; Lecoq et al., 2016; Seo et al., 2015; van der Vlugt et al., 2015, Adams et al., 2014 Verbeek et al., 2014; Batuman et al., 2010; Verbeek et al., 2010; Verbeek et al., 2008; Verbeek et al., 2007), and cassava (Carvajal-Yepes et al., 2014). In the frame of a screening process of plant material to investigate the occurrence and diversity of cassava torrado-like viruses from the major producing areas of Brazil, several samples from the Embrapa - National Cassava and Fruits Research Center collection and Embrapa Amazônia Oriental were analyzed by HTS for virus discovery. The sequence analysis revealed that the assembled sequenced grouped in different clusters: a set of CsTLV sequences clustered with the formerly reported reference sequence from Colombia (Carvajal-Yepes et al., 2014) and another set was determined to be a new cassava torrado-like virus species according to the species demarcation criteria in the family *Secoviridae*, for which we propose the name CsTLV-2. This sequence shared less than 66.3% amino acid identity in the Pro-Pol region and less than 66.4% in the combined CP regions with the CsTLV originally found in Colombia.

Torrado-like virus sequences were found in many of the analyzed samples showing that these viruses are widespread in cassava in Brazil. Although mixed infections in cassava have been described among viruses belonging to different species (Oliveira et al., 2020; Carvajal-Yepes et al., 2014), mixed infections of CsTLV-1 and CsTLV-2 were not detected in this work. The CsTLV sequence assembled from the GenBank dataset originated from a Colombian sample was closer to the Brazilian sequences than to the formerly reported Colombian sequence, showing that sequence diversity does not correlate with geographic location. Comparisons between sequence isolates of CsTLV-2 from different regions of Brazil showed that CsTLV-2 is genetically diverse in Brazil.

Even though the genome organization of torradoviruses is conserved among species, the RNA2 encoded polyprotein processing strategy and cleavage sites have not been reported for all torradoviruses. The reported cleavage sites of the RNA2 polyprotein are Q/A, Q/A and Q/S for ToMarV; Q/V, Q/S and Q/V for ToChSV and Q/S, Q/G and Q/L for Carrot torradovirus for regions between the putative MP/Vp35(Vp37) proteins, Vp35/Vp26(Vp27) and Vp26/Vp24(Vp26) coat proteins, respectively (Ferriol et al., 2016; Gaafar & Ziebell, 2019). Although these sites may vary depending on the viral species, the glutamine (Q) is conserved at the cleavage site in all torradoviruses. The molecular mass of the peptides is also variable (Ferriol et al., 2016), and based on alignment prediction the putatively calculated MW of CP-A, CP-B and CP-C of CsTLV varied from 23.9 to 26 kDa (predicted cleavage sites Q/G, Q/A and Q/V), and the MW of MP from 52.5 to 55.3 kDa. Further experimental studies should be conducted to verify the predicted cleavage sites of torrado-like viruses infecting cassava, especially because all polyproteins encoded by CsTLV RNA1 showed an isoleucine or glutamine instead of a histidine

residue at the substrate binding site (Bazan and Fletterick, 1988). Moreover, some nepoviruses and sequi-viruses do have leucine or cysteine as binding pockets (Sanfaçon et al., 2009).

The main disease that affects cassava in South America is the CFSD. In Brazil, the CsTLV was recently detected in cassava samples with typical symptoms of frogskin disease that tested negative for phytoplasmas (Oliveira et al., 2020). Investigation on the sanitary status of CFSD material is critical to get insights into the etiology of the disease. In this context, the developed RT-PCR tests can be implemented in screening processes to test for the presence of the virus in leaf or tuber material.

In conclusion, our results show the presence of at least two distinct torradovirus species in cassava in Brazil and we propose these viruses be named cassava torrado-like virus 1 (CsTLV-1) and cassava torrado-like virus 2 (CsTLV-2). This work contributes to the understanding of sequence diversity among CsTLV sequences from South America.

4. Materials and Methods

4.1. Plant material

Germplasm accessions received from the collection of Embrapa - National Cassava and Fruits Research Center were maintained and propagated *in vitro* at the Plant Virus Department of the Leibniz Institute DSMZ. Plantlets were transferred to soil, acclimated and a growing-on test was conducted for a period of 6-8 months in glasshouse conditions. In addition, samples collected in trial fields of Embrapa Amazônia Oriental (Belém, Pará, Brazil) (Table S1) were included in the analysis. Single or pooled samples of leaves or tubers were considered for HTS and processed as described below. Leaves from in vitro propagation material or plants maintained in the greenhouse were used for testing by RT-PCR assays and Sanger sequencing.

4.2. Library preparation and high-throughput sequencing

Total RNA was extracted from leaves or tubers using a phenol/chloroform extraction protocol or the GenCatch™ Total RNA Extraction System (Epoch). Complementary DNAs was synthesized using random octamer primers and used as input in the NEBNext Ultra II non-directional RNA Second-Strand Synthesis Module kit (New England Biolabs) for second-strand cDNA synthesis. Twenty-eight libraries were prepared following the NexteraXT DNA Library Prep Kit (Illumina). Following quality control (Bioanalyzer, Agilent) and quantification (Qbit Fluoremeter, Thermo-Fisher Scientific), the libraries were sequenced on MiSeq or NextSeq-500 sequencing platforms at the Leibniz Institute-DSMZ, as paired reads 2x301 nt and 2x151 nt, respectively. For the sample

from Amazon, RNA was isolated from a semi-purified preparation from cassava leaves using the ZR Plant RNA Miniprep Kit (Zymo Research, Irvine, USA) following the manufacturer's protocol. The RNA was sequenced using the Illumina HiSeq 2500 platform with 100-base paired-end reads at the Centro de Genômica Funcional-ESALQ (University of São Paulo, Piracicaba, Brazil), totaling 29 HTS of samples from Brazil

4.3. Bioinformatics analysis of HTS datasets

Beside the 29 datasets generated in this work, an SRA dataset available from the National Center for Biotechnology Information (NCBI; Bioproject: PRJNA231851, Run: SRR1050897), obtained by sequencing of cassava material from CIAT germplasm (Colombia), was also included in our analyses. The bioinformatics workflow followed a pipeline previously described (Blawid et al., 2017). Briefly, read quality was checked with FastQC v. 0.11.9 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and trimming was performed with Trimmomatic v. 0.36. (Bolger et al., 2014) using the following parameters: ILLUMINACLIP: NEXTERAPE-PE.FA: 2:30:10, LEADING: 20, TRAILING: 20, SLIDINGWINDOW: 4:20 AND MINLEN: 36. After trimming, reads were de novo assembled with SPAdes v3.11.1 (Bankevich et al., 2012) using k-values of 21, 33, 55, 77 and 99. Contigs alignments (tBlastx) were calculated in Geneious Prime v. 11.0.5 using a custom plant virus database constructed from available sequences at NCBI. To obtain complete virus genomic sequences, reads were iteratively mapped to the assembled contigs. Because of gaps in the assembled sequences of sample AJB1_86, the virus genomes excluding the terminal ends was completed by Sanger sequencing. The complete coding sequence of CsTLV-2 RNA1 from Amazon were closed with the following primer pairs: torrado1-1 (5'-AAT TTG TGA GTC AGC CAA CGT G-3'), torrado-2 (5'-AGC ATC CTG ACA GCC AAC CTA A-3'), and torrado-1-3 (5'-GTT GAA GTG AAG TAC TTG TAC GCT-3') / torrado-1-4 (5'-GAC TCA AAT GCT GAT GTT CTA ATC CA-3'). The missing fragment in RNA2 was amplified with primers: torrado-2-5 (5'-TGA TCA TGA TCT GGG CTC TGC G-3') and torrado-2-6 (5´-CCA CAA GGC GCT AGT AAT GT -3'). The cDNA synthesis was performed with the SuperScript IV Reverse Transcriptase (Invitrogen™) using a N6 random primer, and used as template for PCR. Fragments were amplified with Phusion High-Fidelity PCR Master Mix (Thermo Scientific™), excised from gel, purified and sequenced by Sanger at the Universidade Federal de Pernambuco. The genomic sequences were annotated in Geneious Prime (Biomatters), using formerly annotated torradovirus sequences as template for comparison. Computational analysis with ScanProsite (De Castro et al., 2006) allowed to search the predicted proteins for conserved domains. Coat protein cleavage sites

were predicted using the carrot torradovirus 1-RNA2 sequence (GenBank acc. NC_025480) as reference.

4.4. Sequence analysis and phylogenetic reconstructions

Multiple sequence alignment calculated with MAFFT v.7 (Katoh *et al.*, 2019). The best model of protein evolution was estimated with ProtTest v.3.4.2 (Darriba *et al*., 2011), using the Bayesian Information Criterion (BIC) (Schwarz, 1978). Phylogenetic reconstructions were performed with MrBayes v.3.2.6 (Ronquist *et al.*, 2012), using the LG+G model with 10.000.000 generations, 1000 sample frequency and 2500 burn-in, on the amino acid region between the CG motif in the protease-domain sequence and the GDD RdRp motif (467 aa in length, based on the CsTLV1 Sec13 sequence, acc. MF449522), according to former phylogenetic analyses of torradovirus sequences (Verbeek *et al*., 2008). The viral species considered in the analysis were: CsTLV1 (MF449522), CsTLV1 (OK040225), tomato chocolate spot virus (NC_013075), tomato marchitez virus (KT756874, MK733734, NC_010987), tomato necrotic dwarf virus (NC_027926), tomato chocolate virus (FJ560489), tomato torrado virus (KJ940975, KM091449, EU563948, KX132808, NC_009013 and MH587229), red clover torradovirus 1 (KY113159), motherwort yellow mottle virus (NC_035218), lettuce necrotic leaf curl virus (NC_035214), carrot torradovirus 1 (LC436363, MN399684, NC_025479). Bayesian phylogenetic reconstruction of RNA2 were performed on the concatenated region including the start codon of the hypothetical protein and the stop codon of the polyprotein (1550 aa in length, based on the CsTLV1 acc. MF449523), using in addition the following sequences: tomato chocolate spot virus (NC_013076), tomato marchitez virus (KT756875, MK726319, NC_010988), tomato necrotic dwarf virus (NC_027927), tomato chocolate virus (FJ560490), tomato torrado virus (KJ940974, KM114266, EU563947, KX132809, NC_009032 and MH587230), red clover torradovirus 1 (KY113160), motherwort yellow mottle virus (NC_035220), lettuce necrotic leaf curl virus (NC_035219), carrot torradovirus 1 (C436364, MN399685, NC_025480), cassava torrado-like virus 1 (MF449523) and CsTLV1 (OK040226). Phylogenetic trees were visualized in Figtree software, v.1.4.4 9 (http://tree.bio.ed.ac.uk/ software/fgtree/). Additionally, the Sequence Demarcation Tool v1.2 (SDT) (Muhire *et al.*, 2014) was employed to calculate identity scores calculated with MAFFT v.7 alignment.

4.5. Detection of CsTLVs by RT-PCR

One-tube RT-PCR assays for detection of CsTLV-1 and CsTLV-2 were developed using primer sets CsTLV-1-F (5´- GCC TCA AGG RAC AGC TTT CT-3´) and CsTLV-1-R (5´-TGC ATG ATG GTT GAT TGC ACA A-3´) and set CsTLV-2-F (5´-CTC TGA GGG TGT TTG TGC TT-

3´) and CsTLV-2-R (5´- GTT TCA TTA TGT TGG TCG CC-3´), to amplify a fragment of ~140bp in the CP-B sequence and of 100 bp in the CP-A sequence, respectively. Total RNA was used as template in the reaction mixture prepared according to the SuperScript™ One-Step RT-PCR System with Platinum™ Taq DNA Polymerase Kit (Invitrogen). The cycling conditions were as follows: 45 ºC for 30 min; 94 ºC for 5 min, 40 cycles of 94 ºC for 30 s, 52 ºC for 30 s and 72 ºC for 20 s; and a final extension at 72 ºC for 2 min. Obtained amplicons were separated by gel-electrophoresis in a 2% agarose gel and visualized under UV light, after staining with ethidium bromide.

Supplementary Materials: Table S1: Characterization of cassava samples, trimming reads and assembly contigs results, Table S2: CsTLV1, CsTLV2 and related viruses: alignment of the RNA1 polyprotein, % id nt and % id aa, Table S3 CsTLV1, CsTLV2 and related viruses: alignment of the RNA2, % id nt and % id aa.

Author Contributions: Conceptualization, G.R.A., P.M., S.W. and R.B; methodology, G.R.A., P.M., S.W. and R.B.; formal analysis, G.R.A, P.M., S.W. and R.B; investigation, G.R.A., P.M., A.J.B., S.W. and R.B.; resources, R.B. and S.W.; data curation, G.R.A., P.M., S.W. and R.B; writing—original draft preparation, G.R.A., P.M., S.W. and R.B; writing—review and editing, G.R.A., P.M., A.J.B., E.J.O., S.W. and R.B.; supervision, R.B. and S.W.; project administration, R.B. and S.W.; funding acquisition, R.B., S.W and E.J.O. All authors have read and agreed to the published version of the manuscript.

Funding: This study was financed in part by the (DSMZ project). The authors also acknowledge the financial support given by the Rural Federal University of Pernambuco (Edital 015/2018-UFRPE PRPG) and by Embrapa Amazônia Oriental

Data Availability Statement: The datasets generated and analyzed during the current study are available from the corresponding author.

Acknowledgments: We thank the Brazilian CAPES foundation for the postgraduate Doctor's degree fellowship to G.R.A, and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for the Research Productivity Fellowship to E.J.O.We also thank the Crop Clinic of Pernambuco from the Rural Federal University of Pernambuco for supporting this work.

This work was supported by Rural Federal University of Pernambuco (Edital 015/2018-UFRPE PRPPG).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Adams IP, Skelton A, Macarthur R, Hodges T, Hinds H, Flint L, Nath PD, Boonham, Fox A (2014) Carrot yellow leaf virus is associated with carrot internal necrosis. PLoS One 9:e109125. doi: 10.1371/journal.pone.0109125
2. Alvarez E, Mejía JF, Llano GA, Loke JB, Calari A, Duduk B, Bertaccini A (2009). Characterization of a phytoplasma associated with frogskin disease in cassava. Plant Dis. 93:1139-1145. https://doi.org/10.1094/PDIS-93-11-1139
3. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko I, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. Journal of Computational Biology 19: 455-477. doi: 10.1089/cmb.2012.0021
4. Bazan JF, Fletterick RJ (1988) Viral cysteine proteases are homologous to the trypsin-like family of serine proteases: structural and functional implications. Proc Natl Acad Sci U S A 21: 7872-7876. doi: 10.1073/pnas.85.21.7872.
5. Blawid R, Silva JMF, Nagata T (2017) Discovering and sequencing new plant viral genomes by next-generation sequencing: description of a practical pipeline. Annals of Applied Biology 170: 301-314. doi: https://doi.org/10.1111/aab.12345
6. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30: 2114-2120. doi:10.1093/bioinformatics/btu170
7. Bruenn JA (2003) A structural and primary sequence comparison of the viral RNA-dependent RNA polymerases. Nucleic Acids Res 31:1821-1829. doi:10.1093/nar/gkg277
8. Carvajal-Yepes M, Olaya C, Lozano I, Cuervo M, Castaño M, Cuellar WJ (2014) Unraveling complex viral infections in cassava (*Manihot esculenta* Crantz) from Colombia. Virus Res 186:76–86. https://doi.org/10.1016/j.virusres.2013.12.011
9. Calvert LA, Cuervo M, Lozano I, Villareal N, Arroyave J (2008). Identification of three strains of a virus associated with cassava plants affected by frogskin disease. Journal of Phytopathology 156: 647–653. https://doi.org/10.1111/j.1439-0434.2008.01412.x

10. Costa AS (1940). Observações sobre o mosaico comum e o mosaico das nervuras da mandioca (*Manihot utilissima* Pohl.) Jornal de Agronomia, 3: 239-248.
11. Darriba D, Taboada GL, Doallo R, Posada D (2011) ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics 8:1164-1165. doi: 10.1093/bioinformatics/btr088.
12. De Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, Bairoch A, Hulo N (2006) ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res 34: 362–365. doi: 10.1093/nar/gkl124.
13. Katoh K, Rozewicki J, Yamada KD (2019) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform 20:1160-1166. doi: 10.1093/bib/bbx108
14. Ferriol I, Silva-Junior DM, Nigg JC, Zamora-Macorra EJ, Falk BW (2016) Identification of the cleavage sites of the RNA2-encoded polyproteins for two members of the genus Torradovirus by N-terminal sequencing of the virion capsid proteins. Virology 496:109–115. 10.1016/j.virol.2016.08.014
15. Gaafar YZA, Ziebell H (2019) Complete genome sequence of a highly divergent carrot torradovirus 1 strain from *Apium graveolens*. Arch Virol 164:1943-1947. doi: 10.1007/s00705-019-04272-3.
16. Hernández A, Calderón H, Zárate RD, Lozano JC (1975) El cuero de sapo de la yuca (*Manihot esculenta*, Crantz). Not Fitopatol 4:117–118.
17. Hillocks R, Raya M, Mtunda K, Kiozia H (2008) Effects of brown streak virus disease on yield and quality of cassava in Tanzania. J. Phytopathol. 149:389-394. https://doi.org/10.1111/j.1439-0434.2001.tb03868.x
18. Lecoq H, Verdin E, Tepfer M, Wipf-Scheibel C, Millot p, Dafalla G, Desbiez C (2016) Characterization and occurrence of squash chlorotic leaf spot virus, a tentative new torradovirus infecting cucurbits in Sudan. Arch Virol 161:1651-1655. doi: https://doi.org/10.1007/s00705-016-2797-8
19. McCallum EJ, Anjanappa RB, Gruissem W (2017) Tackling agriculturally relevant diseases in the staple crop cassava (*Manihot esculenta*). Current Opinion in Plant Biology 38: 50-58. https://doi.org/10.1016/j.pbi.2017.04.008
20. Muhire BM, Varsani A, Martin DP (2014) SDT: A Virus Classification Tool Based on Pairwise Sequence Alignment and Identity Calculation. PLoS ONE 9:e108277. doi:10.1371/journal.pone.0108277

21. Oliveira SAS, Ferreira CF, Diamantino MSAS, Santos TA, Pereira JD, Oliveira EJ (2020) First report of cassava torrado-like virus, cassava polero-like virus and cassava new alphaflexivirus associated with cassava frogskin disease in Brazil. Journal of Plant Pathology. 10: 247. https://doi.org/10.1007/s42161-019-00384-6
22. Owor B, Legg J P, Okao-Okuja G, Obonyo R, Ogenga-Latigo MW (2004) The effect of cassava mosaic geminiviruses on symptom severity, growth and root yield of a cassava mosaic virus disease-susceptible cultivar in Uganda. Annals of Applied Biology, 145: 331-337. doi: 10.1111/j.1744-7348.2004.tb00390.x
23. Pineda B, Jayasinghe U, Lozano JC (1983) La enfermedad "Cuero de Sapo" en yuca (*Manihot esculenta*, Crantz). ASIAVA 4:10-12.
24. Posada D (2008) jModelTest: phylogenetic model averaging. Mol Biol Evol 7:1253-1256. doi: 10.1093/molbev/msn083.
25. Poch O, Sauvaget I, Delarue M, Tordo N (1989) Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. EMBO J 8:3867-74.
26. Sanfaçon H, Wellink J, Le Gall O, Karasev A, van der Vlugt R, Wetzel T (2009) Secoviridae: a proposed family of plant viruses within the order Picornavirales that combines the families Sequiviridae and Comoviridae, the unassigned genera Cheravirus and Sadwavirus, and the proposed genus Torradovirus. Arch Virol 5:899-907. doi: 10.1007/s00705-009-0367-z.
27. Silva DM, Kitajima EW, Oliveira AR (1963) Obtenção do vírus do mosaico comum da mandioca purificado. Ciência e Cultura 15: 304.
28. Seo JK, Kang M, Kwak HR, Kim MK, Kim CS, Lee SH, Kim JS, Choi HS (2015) Complete genome sequence of motherwort yellow mottle virus, a novel putative member of the genus Torradovirus. Arch Virol 160:587–590. doi: https://doi.org/10.1007/s00705-014-2254-5
29. 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: efcient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61:539-542. https://doi.org/10.1093/sysbio/sys029
30. Van der Vlugt RA, Verbeek M, Dullemans AM, Wintermantel WM, Cuellar WJ, Fox A, Thompson JR (2015) Torradoviruses. Annu Rev Phytopathol 53:485-512. doi: 10.1146/annurev-phyto-080614-120021.

31. Verbeek M, Dullemans AM, van den Heuvel JF, Maris PC, van der Vlugt RA (2007) Identification and characterisation of tomato torrado virus, a new plant picorna-like virus from tomato. Arch Virol 152: 881-890. doi:10.1007/s00705-006-0917-6
32. Verbeek M, Dullemans A, van den Heuvel J, Maris PC, van der Vlugt RAA (2008) Tomato marchitez virus, a new plant picorna-like virus from tomato related to tomato torrado virus. Arch Virol 153: 127-134. doi: https://doi.org/10.1007/s00705-007-1076-0
33. Verbeek M, Dullemans AM, van Raaij HM, Verhoeven, JT, van der Vlugt RA (2014) Lettuce necrotic leaf curl virus, a new plant virus infecting lettuce and a proposed member of the genus Torradovirus. Arch Virol 159: 801-805. doi: 10.1007/s00705-013-1835-z
34. Verbeek M, Dullemans A, van den Heuvel H, Maris P, van der Vlugt R. (2010) Tomato chocolàte virus: a new plant virus infecting tomato and a proposed member of the genus Torradovirus. Arch Virol 155: 751-755. doi: 10.1007/s00705-010-0640-1.
35. Walter B, Ladeveze I, Etienne L, Fuchs M (1989). Some properties of a previously undescribed virus from cassava: Cassava American Latent Virus. Annals of Applied Biology 115: 279–289. ttps://doi.org/10.1111/j.1744-7348.1989.tb03386.x
36. Wintermantel WM, Hladky LL, Cortez, AA (2018). Genome sequence, host range, and whitefly transmission of the torradovirus Tomato necrotic dwarf virus. Acta Horticulturae 1207: 295–302. doi:10.17660/actahortic.2018.1207.41
37. Zinga I, Chiroleu F, Legg J, Lefeuvre P, Komba EK, Semballa S, Yandia SP, Mandakombo NB, Reynaud B, Lett J (2013) Epidemiological assessment of cassava mosaic disease in Central African Republic reveals the importance of mixed viral infection and poor health of plant cuttings. Crop Prot 44: 6-12. https://doi.org/10.1016/j.cropro.2012.10.01

Table S1. Overview of the libraries, generated HTS datasets and bioinformatics analysis.

Sample	Denomination	Source	Platform	Sample type	Reads before trimming	Reads after trimming	% Reads trimmed	Contigs generated	Contigs assigned to CsTLV	Reads mapping RNA1	Reads mapping RNA2
1	AJB1_S86	Amazon*	HiSeq 25000	Leaf	17254734	15147828	12.2	56524	42	1110	346
2	DSC369	growing-on test	MiSeq	Leaf	1164386	1143632	1,8	4135	8	454	739
3	DSC372	growing-on test	MiSeq	Leaf	1480066	1458458	1,5	11692	0	0	0
4	DSC382	growing-on test	MiSeq	Leaf	2033140	2007132	1,3	13513	0	0	0
5	DSC389	growing-on test	MiSeq	Leaf	1448456	1261554	12,9	8271	0	0	0
6	DSC404	growing-on test	MiSeq	Leaf	932070	791224	15,1	4160	0	0	0
7	DSC407	growing-on test	MiSeq	Leaf	4827040	4650694	3,7	48033	9	2808	3390
8	DSC409	growing-on test	MiSeq	Leaf	1294094	1278450	1,2	10237	7	395	730
9	DSC416	growing-on test	MiSeq	Leaf	2767482	2724396	1,6	42359	30	334	71
10	DSC417	growing-on test	MiSeq	Leaf	1516120	1444600	4,7	4677	0	0	0
11	DSC430	growing-on test	MiSeq	Leaf	2073210	2055348	0,9	4524	0	0	0
12	DSC435	growing-on test	MiSeq	Leaf	362290	329982	8,9	5825	0	0	0
13	DSC437	growing-on test	MiSeq	Leaf	797534	788134	1,2	4766	9	56	183
14	DSC442	growing-on test	MiSeq	Leaf	1366734	1352452	1,0	12399	0	0	0
15	DSC445	growing-on test	MiSeq	Leaf	2517382	2487282	1,2	23990	0	0	0

16	DSC492	growing-on test	MiSeq	Leaf	1686030	1651836	2,0	13529	12	914	1157
17	DSCG1	growing-on test	MiSeq	Leaf	424616	399974	5,8	5832	0	0	0
18	DSCG12	growing-on test	MiSeq	Leaf	467798	433786	7,3	5731	13	12	22
19	DSCG2	growing-on test	MiSeq	Leaf	438500	418144	4,6	10372	30	329	443
20	DSCG3	growing-on test	MiSeq	Leaf	431404	412506	4,4	11107	0	0	0
21	DSCG4	growing-on test	MiSeq	Leaf	395616	374710	5,3	9340	12	840	1173
22	DSCG5	growing-on test	MiSeq	Leaf	2924524	2595720	11,2	36905	37	837	1530
23	DSCG6	growing-on test	MiSeq	Leaf	510208	507104	0,6	11835	0	0	0
24	DSCG7	growing-on test	MiSeq	Leaf	3604914	3479942	3,5	47488	12	401	688
25	P25_6A	growing-on test	NextSeq	Tuber	13498456	12201204	9,6	58645	0	0	0
26	P26_P27_4C	growing-on test	MiSeq/NextSeq	Leaf	98271182	86425598	12,1	253292	44	11691	20488
27	P26-2E**	growing-on test	NextSeq	Leaf	6656458	6167118	7,4	51691	11	63590	121343
28	P27_6F	field sample	MiSeq	Leaf	48781540	45816820	6,1	89672	24	23621	25367
29	P27_6G	field sample	NextSeq	Leaf	66773560	61845420	7,4	91373	26	74113	80833
30	SRR1050897	CIAT germplasm collection	Genome Analyzer II	Stem	13725068	11253093	18,0	43413	19	6059	28337

* Embrapa Amazônia Oriental

**DSMZ isolate PV-1279 from Embrapa Mandioca e Fruticultura, Cruz das Almas, Bahia, Brazil.

Table S2. Percentages of identity of the RNA1 polyprotein (nt and aa) of CsTLV-1, CsTLV-2 and other torradoviruses.

RNA1 (nt)	CaTV1 NC_025479	CsTLV-1 DSC407	CsTLV-1 DSC409	CsTLV-1 DSC492	CsTLV-1 DSCG4	CsTLV-1 DSCG5	CsTLV-1 P26_2E	CsTLV-1 SRR1050897	CsTLV-1 MF449522	CsTLV-1 OK040225	CsTLV-2 AJ1_S86	CsTLV-2 P26_P27_4C	CsTLV-2 P27_6F_51	CsTLV-2 P27_6G_1	CsTLV-2 P27_6G_2	LNLCV NC_035214	MyMoV NC_035218	ReCTV1 KY113159	ToChSV NC_013075	ToChV FJ560489	ToMarV NC_010987	ToNDV NC_027926	ToTV NC_009013
NC_025479 carrot torradovirus 1	100,0																						
DSC407 cassava torrado-like virus 1	53,4	100,0																					
DSC409 cassava torrado-like virus 1	55,4	96,5	100,0																				
DSC492 cassava torrado-like virus 1	53,7	95,7	95,4	100,0																			
DSCG4 cassava torrado-like virus 1	53,1	98,4	96,5	95,8	100,0																		
DSCG5 cassava torrado-like virus 1	54,0	98,2	96,5	95,8	98,4	100,0																	
P26_2E cassava torrado-like virus 1	53,8	95,6	95,3	99,9	95,7	95,7	100,0																
SRR1050897 cassava torrado-like virus 1	53,6	97,2	97,3	96,2	97,2	97,2	96,1	100,0															
MF449522 cassava torrado-like virus 1 Sec13	53,8	87,8	87,2	88,0	87,9	87,7	87,9	88,1	100,0														
OK040225 cassava Torrado-like virus 1 Yop12	35,7	61,0	61,7	60,9	61,0	60,9	60,8	60,9	68,7	100,0													
AJ1_S86 cassava Torrado-like virus 2	54,0	57,1	57,2	57,4	57,3	57,2	57,4	57,5	57,4	37,9	100,0												
P26_P27_4C cassava torrado-like virus 2	53,9	57,7	57,2	57,4	57,7	57,4	57,4	57,5	57,6	38,5	83,3	100,0											
P27_6F cassava torrado-like virus 2	54,2	58,4	57,9	58,3	58,2	58,3	58,3	58,6	58,2	38,1	85,3	83,1	100,0										
P27_6G_1 cassava torrado-like virus 2	54,1	57,7	57,8	57,8	57,9	57,8	57,9	58,1	58,0	38,2	84,7	88,5	94,3	100,0									
P27_6G_2 cassava torrado-like virus 2	54,0	57,7	57,3	57,5	57,7	57,7	57,5	57,6	57,5	38,4	83,5	97,9	84,8	90,4	100,0								
NC_035214 lettuce necrotic leaf curl virus	57,7	53,0	53,8	53,8	52,6	53,2	53,9	53,1	53,2	36,3	52,7	52,2	52,5	52,4	52,5	100,0							
NC_035218 motherwort yellow mottle virus	58,8	53,6	53,5	52,7	52,9	53,4	53,3	53,2	53,6	35,4	52,7	52,8	52,6	53,1	53,0	58,8	100,0						
KY113159 red clover torradovirus 1	58,9	53,5	54,9	53,6	53,8	54,2	53,6	53,8	54,3	34,9	54,1	53,3	53,3	53,1	53,3	58,3	59,3	100,0					
NC_013075 tomato chocolate spot virus	54,6	53,4	53,6	53,3	53,3	53,6	53,4	53,5	53,8	35,6	53,9	53,7	54,0	53,0	53,3	54,1	53,6	54,4	100,0				
FJ560489 tomato chocolate virus	55,6	54,2	54,1	54,2	54,5	54,5	54,2	53,7	53,8	35,3	53,7	54,0	53,2	53,3	53,5	53,6	53,4	54,2	67,9	100,0			
NC_010987 tomato marchitez virus	55,0	54,7	55,4	54,0	54,3	54,9	54,0	54,5	54,2	36,3	53,9	53,6	53,5	54,0	53,2	54,2	53,0	54,3	73,6	68,7	100,0		
NC_027926 tomato necrotic dwarf virus	54,8	53,6	54,1	53,4	53,6	53,7	53,3	53,5	53,0	35,4	53,6	52,8	53,6	53,5	53,3	53,8	53,3	53,2	72,0	68,9	73,1	100,0	
NC_009013 tomato torrado virus	54,5	53,5	54,8	53,5	53,8	53,9	53,5	53,5	53,8	35,9	53,6	53,8	53,2	53,3	53,8	53,1	54,6	54,0	64,5	64,9	65,3	64,6	100,0

RNA1 (aa)	CaTV1 NC_025479	CsTLV-1 DSC407	CsTLV-1 DSC409	CsTLV-1 DSC492	CsTLV-1 DSCG4	CsTLV-1 DSCG5	CsTLV-1 P26_2E	CsTLV-1 SRR1050897	CsTLV-1 MF449522	CsTLV-1 OK040225	CsTLV-2 AJ1_S86	CsTLV-2 P26_P27_4C	CsTLV-2 P27_6F_51	CsTLV-2 P27_6G_1	CsTLV-2 P27_6G_2	LNLCV NC_035214	MyMoV NC_035218	ReCTV1 KY113159	ToChSV NC_013075	ToChV FJ560489	ToMarV NC_010987	ToNDV NC_027926	ToTV NC_009013
NC_025479 carrot torradovirus 1	100,0																						
DSC407 cassava torrado-like virus 1	36,2	100,0																					
DSC409 cassava torrado-like virus 1	38,0	99,3	100,0																				
DSC492 cassava torrado-like virus 1	36,3	99,4	99,2	100,0																			
DSCG4 cassava torrado-like virus 1	36,2	99,7	99,4	99,4	100,0																		
DSCG5 cassava torrado-like virus 1	36,1	99,6	99,3	99,4	99,7	100,0																	
P26_2E cassava torrado-like virus 1	36,3	99,3	99,0	99,9	99,3	99,2	100,0																
SRR1050897 cassava torrado-like virus 1	36,2	99,5	99,4	99,4	99,7	99,5	99,2	100,0															
MF449522 cassava torrado-like virus 1 Sec13	36,5	97,3	97,1	97,3	97,3	97,3	97,2	97,3	100,0														
OK040225 cassava Torrado-like virus 1 Yop12	36,4	97,3	97,1	97,3	97,3	97,3	97,2	97,3	99,7	100,0													
AJ1_S86 cassava Torrado-like virus 2	36,8	53,2	53,7	53,3	53,3	53,3	53,3	53,3	53,5	53,5	100,0												
P26_P27_4C cassava torrado-like virus 2	36,7	53,5	53,9	53,5	53,6	53,5	53,5	53,5	53,7	53,6	96,2	100,0											
P27_6F cassava torrado-like virus 2	36,8	53,1	53,7	53,2	53,3	53,1	53,2	53,2	53,4	53,4	97,4	96,2	100,0										
P27_6G_1 cassava torrado-like virus 2	36,8	53,1	53,7	53,2	53,2	53,1	53,2	53,2	53,4	53,4	96,8	97,3	98,8	100,0									
P27_6G_2 cassava torrado-like virus 2	36,8	53,5	53,9	53,5	53,6	53,5	53,5	53,5	53,7	53,6	96,3	99,3	96,6	97,8	100,0								
NC_035214 lettuce necrotic leaf curl virus	53,9	35,4	37,3	35,4	35,5	35,4	35,4	35,5	35,7	35,6	36,2	36,4	36,4	36,4	36,3	100,0							
NC_035218 motherwort yellow mottle virus	53,1	35,2	37,1	35,4	35,3	35,4	35,4	35,3	35,6	35,5	36,4	36,6	36,9	36,9	36,7	53,7	100,0						
KY113159 red clover torradovirus 1	53,8	35,5	37,9	35,5	35,6	35,4	35,5	35,6	35,8	35,8	37,3	37,3	37,2	37,0	37,2	56,3	54,2	100,0					
NC_013075 tomato chocolate spot virus	39,2	35,8	37,2	36,1	36,0	35,8	36,2	36,0	36,4	36,3	36,5	36,9	36,5	36,7	36,9	37,7	37,3	39,0	100,0				
FJ560489 tomato chocolate virus	39,9	36,3	37,5	36,2	36,2	36,0	36,2	35,1	36,1	36,1	36,8	36,5	36,4	36,6	36,5	38,3	37,0	38,4	73,3	100,0			
NC_010987 tomato marchitez virus	40,5	35,9	37,3	36,1	36,0	35,9	36,2	36,0	36,1	36,0	36,2	36,4	36,4	36,7	36,4	38,1	37,5	39,7	83,4	74,6	100,0		
NC_027926 tomato necrotic dwarf virus	40,5	35,7	36,9	35,8	35,7	35,7	35,9	35,7	36,1	36,0	36,8	36,9	36,8	37,1	37,1	38,4	37,9	39,7	81,3	74,0	82,9	100,0	
NC_009013 tomato torrado virus	38,6	34,2	35,2	34,4	34,4	34,4	34,5	34,7	35,1	35,0	36,6	36,5	36,3	36,5	36,4	38,1	36,8	37,6	65,0	64,3	65,3	64,2	100,0

Table S3. Percentages of identity of the RNA2 at nucleotide and concatenated proteins sequences of CsTLV-1, CsTLV-2 and other torradoviruses.

RNA2 (nt)	CaTV1 NC_025480	CsTLV-1 DSC407	CsTLV-1 DSC409	CsTLV-1 DSC492	CsTLV-1 DSCG4	CsTLV-1 DSCG5	CsTLV-1 P26_2E	CsTLV-1 SRR1050897	CsTLV-1 MF449522	CsTLV-1 OK040226	CsTLV-2 AJ1_S86	CsTLV-2 P26_P27_4C	CsTLV-2 P27_6F	CsTLV-2 P27_6G_1	CsTLV-2 P27_6G_2	LNLCV NC_035219	MyMoV NC_035220	ReCTV1 KY113160	ToChSV NC_013076	ToChV FJ560490	ToMarV NC_010988	ToNDV NC_027927	ToTV NC_009032
NC_025480 carrot torradovirus 1	100,0																						
DSC407 cassava torrado-like virus 1	52,8	100,0																					
DSC409 cassava torrado-like virus 1	52,9	99,0	100,0																				
DSC492 cassava torrado-like virus 1	52,7	97,3	97,3	100,0																			
DSCG4 cassava torrado-like virus 1	53,2	99,0	98,7	97,1	100,0																		
DSCG5 cassava torrado-like virus 1	53,0	99,1	98,8	97,2	98,8	100,0																	
P26_2E cassava torrado-like virus 1	52,6	97,3	97,3	99,9	97,1	97,1	100,0																
SRR1050897 cassava torrado-like virus 1	52,8	98,4	98,3	97,5	98,2	98,3	97,4	100,0															
MF449522 cassava torrado-like virus 1 Sec13	52,0	91,4	91,3	91,6	91,5	91,3	91,6	91,4	100,0														
OK040226 cassava Torrado-like virus 1 Yop12	52,7	98,0	97,9	97,2	97,9	98,0	97,2	99,2	91,3	100,0													
AJ1_S86 cassava Torrado-like virus 2	53,3	61,3	61,2	61,5	61,5	61,5	61,5	61,7	60,4	61,7	100,0												
P26_P27_4C cassava torrado-like virus 2	52,2	60,9	61,0	60,7	61,0	60,8	60,7	60,9	61,0	60,9	88,2	100,0											
P27_6F cassava torrado-like virus 2	52,2	60,9	61,0	60,7	61,0	60,8	60,7	60,9	61,0	60,9	88,2	99,9	100,0										
P27_6G_1 cassava torrado-like virus 2	52,4	61,2	61,4	61,5	61,4	61,2	61,5	61,3	61,6	61,1	86,2	92,9	92,8	100,0									
P27_6G_2 cassava torrado-like virus 2	52,1	60,8	61,0	60,6	60,9	60,8	60,6	60,9	61,0	60,9	88,2	99,9	99,8	92,8	100,0								
NC_035219 lettuce necrotic leaf curl virus	55,3	50,8	50,7	50,9	50,5	50,6	50,9	50,6	51,5	51,4	53,2	52,9	53,0	53,2	52,9	100,0							
NC_035220 motherwort yellow mottle virus	57,0	52,1	52,0	52,4	52,5	52,1	52,4	52,4	51,7	52,3	54,3	53,0	52,9	52,9	52,9	57,9	100,0						
KY113160 red clover torradovirus 1	56,8	52,7	52,2	52,6	52,4	52,6	52,2	52,3	51,7	52,1	52,6	51,9	51,8	52,4	51,8	59,5	59,1	100,0					
NC_013076 tomato chocolate spot virus	53,6	52,4	52,0	52,5	52,8	52,3	52,4	52,0	52,1	52,3	53,0	53,0	52,8	53,0	53,0	50,9	51,3	52,3	100,0				
FJ560490 tomato chocolate virus	52,5	53,3	53,5	53,6	53,6	53,0	53,6	54,0	53,7	53,8	52,7	52,5	52,6	52,3	52,5	52,0	52,0	52,2	70,1	100,0			
NC_010988 tomato marchitez virus	51,5	52,4	52,6	53,1	52,4	52,2	53,3	52,5	52,4	52,8	52,4	52,7	52,7	52,6	52,7	51,3	52,6	52,3	73,3	69,5	100,0		
NC_027927 tomato necrotic dwarf virus	54,2	53,2	52,4	52,5	53,0	52,4	52,5	53,0	52,3	53,2	52,5	52,8	52,7	52,0	52,8	51,6	51,8	53,3	73,6	69,8	77,2	100,0	
NC_009032 tomato torrado virus	52,1	51,1	50,1	50,7	50,7	50,6	50,7	50,2	49,6	50,7	52,5	52,5	52,5	52,5	52,5	51,5	50,4	50,6	66,4	64,6	65,9	65,9	100,0

RNA2 (aa)	CaTV1 NC_025480	CsTLV-1 DSC407	CsTLV-1 DSC409	CsTLV-1 DSC492	CsTLV-1 DSCG4	CsTLV-1 DSCG5	CsTLV-1 P26_2E_S13	CsTLV-1 SRR1050897	CsTLV-1 MF449522	CsTLV-1 OK040226	CsTLV-2 AJ1_S86	CsTLV-2 P26_P27_4C	CsTLV-2 P27_6F	CsTLV-2 P27_6G_1	CsTLV-2 P27_6G_2	LNLCV NC_035219	MyMoV NC_035220	ReCTV1 KY113160	ToChSV NC_013076	ToChV FJ560490	ToMarV NC_010988	ToNDV NC_027927	ToTV NC_009032
NC_025480 carrot torradovirus 1	100,0																						
DSC407 cassava torrado-like virus 1	38,0	100,0																					
DSC409 cassava torrado-like virus 1	38,0	99,9	100,0																				
DSC492 cassava torrado-like virus 1	38,2	99,4	99,4	100,0																			
DSCG4 cassava torrado-like virus 1	38,0	99,9	99,9	99,5	100,0																		
DSCG5 cassava torrado-like virus 1	38,0	99,9	99,9	99,5	100,0	100,0																	
P26_2E cassava torrado-like virus 1	38,2	99,4	99,4	100,0	99,5	99,5	100,0																
SRR1050897 cassava torrado-like virus 1	38,1	99,6	99,6	99,5	99,7	99,7	99,5	100,0															
MF449522 cassava torrado-like virus 1 Sec13	37,9	98,3	98,2	98,4	98,4	98,4	98,4	98,4	100,0														
OK040226 cassava Torrado-like virus 1 Yop12	38,0	99,6	99,6	99,5	99,7	99,7	99,5	99,7	98,4	100,0													
AJ1_S86 cassava Torrado-like virus 2	38,4	60,7	60,7	60,7	60,7	60,7	60,7	60,9	60,7	60,7	100,0												
P26_P27_4C cassava torrado-like virus 2	37,9	61,1	61,1	61,1	61,1	61,1	61,1	61,2	60,8	61,1	98,2	100,0											
P27_6F cassava torrado-like virus 2	37,9	61,1	61,1	61,1	61,1	61,1	61,1	61,2	60,8	61,1	98,2	100,0	100,0										
P27_6G_1 cassava torrado-like virus 2	38,6	60,6	60,6	60,6	60,6	60,6	60,6	60,8	60,6	60,7	97,3	97,6	97,6	100,0									
P27_6G_2 cassava torrado-like virus 2	37,9	61,1	61,1	61,1	61,1	61,1	61,1	61,2	60,8	61,1	98,2	100,0	100,0	97,6	100,0								
NC_035219 lettuce necrotic leaf curl virus	51,1	35,7	35,7	35,8	35,7	35,7	35,8	35,8	35,7	35,7	35,6	35,5	35,5	35,5	35,5	100,0							
NC_035220 motherwort yellow mottle virus	53,4	36,6	37,1	36,8	36,6	36,6	36,8	36,8	36,9	36,8	37,7	37,4	37,4	37,4	37,4	55,0	100,0						
KY113160 red clover torradovirus 1	52,2	36,2	36,2	36,3	36,2	36,2	36,3	36,4	36,3	36,3	37,7	37,2	37,2	37,2	37,2	57,9	58,0	100,0					
NC_013076 tomato chocolate spot virus	37,4	34,1	34,2	34,3	34,1	34,1	34,3	34,2	35,2	34,2	36,1	36,2	36,2	36,4	36,2	35,8	37,1	36,7	100,0				
FJ560490 tomato chocolate virus	36,0	35,4	35,6	35,3	35,4	35,4	35,3	35,5	35,2	35,5	35,6	35,7	35,7	36,6	35,7	35,0	36,8	36,0	76,9	100,0			
NC_010988 tomato marchitez virus	37,4	34,6	34,7	34,7	34,6	34,6	34,7	34,6	34,7	34,7	34,6	34,4	34,4	34,0	34,4	36,5	37,1	36,0	83,0	76,4	100,0		
NC_027927 tomato necrotic dwarf virus	38,1	35,0	35,1	34,7	35,0	35,0	34,7	34,7	35,3	34,7	35,3	34,6	34,6	34,6	34,6	36,4	37,8	36,8	83,7	76,0	89,7	100,0	
NC_009032 tomato torrado virus	37,3	33,1	33,1	33,4	33,1	33,1	33,4	33,1	33,3	33,1	35,7	35,6	35,6	35,5	35,6	35,6	36,6	34,8	69,6	68,9	69,4	69,1	100,0

CAPÍTULO III

Molecular characterization of cassava virus X in Brazil

Tropical Plant Pathology

Molecular characterization of cassava virus X in Brazil

[1]Géssyka Rodrigues de Albuquerque, [1]José Ailton Cruz Macedo dos Santos, [2]Alessandra de Jesus Boari, [2]Elisa Ferreira Moura Cunha, [2]Késsia Fátima Cunha Pantoja, [3]Stephan Winter & [1]Rosana Blawid

1 Universidade Federal Rural de Pernambuco, Recife, Pernambuco, Brazil;
2 Embrapa Amazônia Oriental, Belém, Pará, Brazil;
3 Plant Virus Department, Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, Germany.

Abstract

Cassava (*Manihot esculenta* Crantz.) is an important crop in tropical regions and viral diseases like cassava mosaic disease caused by the potexvirus cassava common mosaic virus (CsCMV) may be responsible for losses up to 30% in cassava production. In addition to CsCMV, other potexviruses have been reported infecting cassava plants, including cassava virus X (CsVX) and cassava colombian symptomless virus (CsCSV) The first report of CsVX in South America was from 1985, and so far, no CsVX infections in Brazil have been described. Here, we identified for the first time CsVX and as well as a satellite virus infecting cassava plants in the Amazon region using High-Throughput Sequencing (HTS) and Sanger sequencing. The Colombian CsVX sequence (SRR1050897) shares 87.2 % nt and 97.2 % aa identity with the ORF1 sequence region of CsVX from Venezuela (NC_034375). Nucleotide and amino acid sequence comparisons of the CP region of CsVX showed the highest identities of 84,7 % and 96,7 % followed by 82,3 % and 97,2 %, respectively, with an assembled CsVX sequence from Colombia (Bioproject PRJNA231851, SRR1050897) and Venezuela (NC_034375). In additional a novel CsVX sequence was found in the SRA data from China (Bioproject PRJNA578024, SRR10480878), suggesting a possible introduction of this virus. For the first time we describe nearly complete sequences of CsVX from Brazil, China and Colombia. The potential impacts of infections in cassava by the new potexvirus and the satellite virus in Brazil are yet to be determined.

Keywords: *Manihot esculenta* (Crantz.); cassava potexviruses; High-Throughput Sequencing

Introduction

Cassava (*Manihot esculenta* Crantz.) is an important crop in tropical regions (McCallum et al. 2017). In Brazil, cassava is widely produced for starch (sweet and sour), flours, tapioca or fresh (in natura) consumption. For the last 10 years Brazil´s production has stood between 22 to 25 million tonnes of roots a year, leading as the fourth largest cassava producer in the world (FAO, 2020). Virus diseases like the cassava common mosaic disease caused by the potexvirus cassava common mosaic virus (CsCMV) are serious impediments to crop productivity and can cause losses up to 30% in the Brazilian cassava production (Venturini et al., 2016). In addition to CsCMV, cassava virus X (CsVX) (Chaparro-Martínez and Trujillo-Pinto 2001), and cassava colombian symptomless virus (CsCSV) have been found in cassava plants (Alabi et al. 2011; Carvajal-Yepes et al. 2014, Oliveira et al. 2020) also occurring in mixed infections, in plants from Colombia showing symptoms of Frog Skin Disease (FSD) (Harrison et al. 1986, Lennon et al. 1985, Lennon et al. 1986, Nolt et al. 1992). Potexviruses are mechanically transmitted viruses and have no known vectors. Because cassava is a vegetatively propagated crop, multiple viruses can accumulate and are maintained over cultivation cycles in infected plants grown from infected cassava stem cuttings (Lozano et al. 2017).

The genus *Potexvirus* in the family *Alphaflexiviridae* has 49 species approved by the International Committee on Taxonomy of Viruses (ICTV). The single-stranded genomic RNA is between 5800-7000n size and encodes five ORFs flanked by a 5′-methyl guanosine cap and a 3′-polyadenylated tail. The ORF1 encodes the viral replicase (RdRp) (King et al. 2012). The overlapping ORFs 2, 3 and 4 encode the triple block proteins, TGB1, TGB2 and TGB3, respectively, that are involved in cell-to-cell t and long-distance movement (Park et al. 2014). The ORF5 encodes the coat protein (CP), which is also required for cell-to-cell movement (Kreuze et al. 2020) and virion assembly. In cassava Colombian symptomless virus and cassava virus X TGB3 is missing.

Previous studies have shown that the viral diversity in cassava is underestimated (Carvajal-Yepes et al. 2014; Zinga et al. 2013). We therefore undertook this study to resolve the diversity of viruses infecting cassava in Amazon, Brazil. Using High-throughput sequencing (HTS), we investigated cassava samples from Amazon, to reveal the presence of potexviruses, which we compared to virus sequences deposited in the Sequence Read Archive (SRA) of NCBI.

Material and methods

RNA extraction, High-Throughput Sequencing, and SRA database analysis

Total RNA was isolated asymptomatic cassava leaves collected from the Cassava Germplasm Bank area of Embrapa Amazônia Oriental (Amazon, Brazil) using the PureLink$^{@}$ kit for viral RNA/DNA Miniprep (Thermo Fisher Scientific/Invitrogen, MA, USA) following the manufacturer's protocol. The RNA was sequenced (AJB1_S86_R1.fastq and AJB1_S86_R2.fastq) using the Illumina HiSeq 2500 platform with 100-base paired-end reads at the Centro de Genômica Funcional-ESALQ (University of São Paulo, Piracicaba, Brazil). In addition, two files available from the Sequence Read Archive (SRA) database of GenBank, Bioproject PRJNA231851-SRR1050897, and Bioproject PRJNA578024-SRR10480878 of Colombia and China, respectively, were used for genome assembly.

Bioinformatic pipeline for assembling viral genomes and secondary structure analysis

The bioinformatics analyses followed a pipeline for plant viral genomes assembly previously described by Blawid et al. 2017. The reads quality was checked using FastQC v. 0.11.9 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and quality and adapter trimming were performed with Trimmomatic v. 0.36. (Bolger et al. 2014), using the following parameters:ILLUMINACLIP:TruSeq3-PE.fa:2:30:10, SLIDINGWINDOW:4:20, MINLEN: 36, LEADING:20 and TRAILING:20. Reads were *de novo* assembled with SPAdes v3.11.1 (Bankevich et al. 2012) in careful mode, using k-values of 21, 33, 55 and 77. After assembling the contigs, tBlastx searchers were performed in Geneious v.11.0.5 using a hand-made plant RefSeq virus database from 9361 sequences downloaded from the NCBI. Viral contigs were extended through consecutive mappings of the reads using the Geneious assembler (performing 5 iterations per mapping with medium-low sensitivity). Genome annotation of viral sequences were performed in Geneious with the help of local alignment tools of the NCBI (https://blast.ncbi.nlm.nih.gov), UniProt (The UniProt Consortium, 2021), Pfam (Mistry et al. 2021) and the Simple Modular Architecture Research Tool (SMART) (Letunic et al. 2021). The final assembled sequence was used for designing specific primers to amplify a fragment of 1355 nt of cassava virus X ORF1. The consensus secondary structures of 5´-UTRs of CsVX and cassava satellite virus were constructed with RNAalifold from the Vienna RNA Web Services (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAalifold.cgi).

Virus detection

Fourteen cassava samples (CPATU 147, CPATU 448, CPATU 452, CPATU 390, CPATU 025, CPATU 016, CPATU 507, CPATU 322, CPATU 232, CPATU 229, CPATU 013, CPATU 250, CPATU 183, and CPATU 592) were collected from the Cassava Germplasm Bank area of Embrapa Amazônia Oriental, Brazil. For virus detection, total RNA was extracted using a modified protocol of the CTAB method (Doyle and Doyle 1990). Two-step RT-PCR was performed using the Moloney murine leukemia virus (MMLV) reverse transcriptase (Invitrogen™) to cDNA synthesis of pooled samples with four RNA each, and oligodeoxyribonucleotides $d(N)_6$ random primers. PCR temperature cycling conditions were as follows: 80 ºC for 1 min for a manually hot start; initial denaturation at 95 ºC for 3 min, 35 cycles of 95 ºC for 30 s, 51 ºC for 1 min and 72 ºC for 1 min and 25 s; and a final extension at 72 ºC for 10 min using the Invitrogen Platinum Taq DNA Polymerase. For this reaction, specific primers potex-624F (5'-CCA AAT AAT GCT CCC AAT TAA GT-3') and potex-1978R (5'-ATA TTG AGT CAG TAG CAC GAG A-3') were employed. RT-PCR of positive pooled samples were tested individually for CsCVX. Amplicons were excised, purified, and sequenced by Sanger sequencing. For the Sanger sequencing of amplified cassava satellite fragments (538 nt), the primers SAT-272F (5'-AAC CAC TCT CGC CAG AAA GG-3') and SAT-809R (5'-CGA TTT GTG GGT CCG GGT AT-3') were used.

Sequence and phylogenetic analysis

Three assembled CsVX sequences (AJB1_S86, SRR1050897 and SRR10480878) and the following sequences of the *Potexvirus* genus available at NCBI were used for phylogenetic analysis: actinidia virus X (AcVX, NC_028649), allium virus X (NC_012211), alstroemeria virus X (AVX, NC_007408), alternanthera mosaic virus (AltMV, AY863024) ambrosia asymptomatic virus 1 (AAV1, KF421905), asparagus virus 3 (AV3, NC_010416), babaco mosaic virus (BabMV, NC_036587), bamboo mosaic virus (BaMV, NC_001642), cactus virus X (CVX, NC_002815), cassava colombian symptomless virus (CsCSV, KC505252), cassava common mosaic virus (CsCMV, KT002439, NC_001658, MW175326, MN243731, MT038420, KT002435 and MN428639), cassava virus X (CsVX, NC_034375), citrus yellow mottle-associated virus (CiYMaV, MK957246), citrus yellow vein clearing virus (CYVCV, NC_026592), cnidium virus X (CnVX, LC460456), cymbidium mosaic virus (CymMV,

NC_001812), euonymus yellow mottle-associated virus (EYMaM, MK572000), euonymus yellow vein virus (EYVV, NC_035190), foxtail mosaic virus (FoMV, NC_001483), hosta virus X (HVX, NC_011544), hydrangea ringspot virus (HdRSV, NC_006943), indian citrus ringspot virus (ICRSV, NC_003093), lettuce virus X (LeVX, NC_010832), lily virus X (LVX, NC_007192), malva mosaic virus (MalMV, NC_008251), mint virus X (MVX, NC_006948), narcissus mosaic virus (NMV, NC_001441), nerine virus X (NVX, NC_007679), opuntia virus X (OpVX, NC_006060), papaya mosaic virus (PapMV, NC_001748), pepino mosaic virus (PepMV, NC_004067), phaius virus X (PhVX, NC_010295), pitaya virus X (PiVX, NC_024458), plantago asiatica mosaic virus (PlAMV, NC_003849), potato aucuba mosaic virus (PAMV, NC_003632), potato virus X (PVX, NC_011620), scallion virus X (SVX, AJ316085), schlumbergera virus X (SchVX, NC_011659), senna mosaic virus (SenMV, NC_030746), strawberry mild yellow edge virus (SMYEV, NC_003794), tamus red mosaic virus (TRMV, NC_016003), tulip virus X (TVX, NC_004322), turtle grass virus X (TGVX, NC_040644), vanilla virus X (VaVX, NC_035205), yam virus X (YVX, NC_025252) and zygocactus virus X (ZyVX, NC_006059).

Multiple sequence alignments were generated with the MAFFT v.7 program (Katoh et al. 2019). The best model of amino acid evolution was estimated with ProtTest v.3.4.2 (Darriba et al., 2011), using the Bayesian Information Criterion (BIC) (Schwarz, 1978). Phylogenetic reconstructions were performed with MrBayes v.3.2.6 (Ronquist et al. 2012), using the LG+I+G+F model for ORF1 and LG+G+F model for CP regions with 10.000.000 generations, 1000 sample frequency and 10 % of burnin. Phylogenetic trees were visualized with Figtree v.1.4.4 (http://tree.bio.ed.ac.uk/software/fgtree/). Dendroscope (version 3.8.2) was used to compare phylogenetic trees using a tanglegram algorithm (Huson and Scornavacca, 2012) and InkScape (https://inkscape.org/) was used to edit tree graphic formats. The identity scores were calculated with the Sequence Demarcation Tool v1.2 (SDT) (Muhire et al. 2014).

Results

High-Throughput Sequencing analysis

Three nearly-complete genomes of CsVX were assembled from the AJB1_S86, SRR1050897 (Colombia) and SRR10480878 (China) datasets. The HTS dataset AJB1_S86 of Amazon generated a total number of viral contigs of 56524, of which 14 were contigs of cassava virus

X. Reads of the SRA database (SRR10480878 and SRR1050897) of GenBank were assembled into 104278 and 43285 contigs, of which 22 and 13 matched to CsVX sequences, respectively (**Table 1**). The near-complete genome sequences of CsVX from Colombia (SRR1050897) were assembled with an average coverage of 146X (1,963 reads); from China (SRR10480878) with 50X (772 reads) and from Brazil with 133X (1,607 reads). Among the contigs that showed high identities with satellite sequences (**Table 1**), the cassava satellite virus sequence (coverage 361X, 1,630 reads) was assembled from the AJB1_S86 data, and it was found only in samples from Brazil.

Table 1 Information of the HTS datasets and bioinformatics analysis.

HTS	Country	RBT*	RAT	%T	TC	Contigs of CsVX
AJB1_S86	Brazil	17254734	15147828	12.2	56524	14
SRR10480878	China	275635662	26210411	9.5	104278	22
SRR1050897	Colombia	13725068	11253093	18	43295	13

*RBT: Reads before trimming, RAT: Reads after trimming, %T: Percentage trimmed, TC: Total contigs

Characterization and genome structure and detection of the cassava virus X and cassava satellite virus

The assembled genomes of CsVX ranged from 5801-5927 nt, excluding Poly-A tails. The full-length ORF1 gene varied in size from 3912 to 3930 nt, and the TGB1 gene from 693 to 699 nt. The TGB2 gene was 300 nt long and the CP was 648 nt in all sequences from Brazil, China and Colombia. The CP sequence from Venezuela consisted of 645 nt (Table 1).

SMART analysis of the coding sequence of CsVX ORF1 revealed three domains of Methyltransferase (Met, aa positions 39-331), Helicase (Hel, aa positions 588-821) and RNA-dependent RNA polymerase (RdRp, aa positions 973-1293) in the order from the N- to C-terminus. The ORF1 aa sequence has the DxxR motif (123-126 aa) and glycine-rich motif GxGxG (208-212 aa) in the Met domain. The GKS (597-599 aa) (Li et al. 2001) and TYSSSQG (778-784 aa) motifs were found in the helicase domain, and the HQQAKDE (917-923 aa), KSQWVKK (1018-1024 aa), GTMARY (1051-1056 aa), QDGAML (1106-1111 aa) (van der VLUG; BERENDSEN, 2002), TGEGPTFDANTE (1151-1162 aa), and GDD (1181-1186 aa) (Lozano et al. 2017) motifs in the RdRp domain. The TGB1 protein contains the AVAGAGKTT (34-42 aa) and PDILDE motifs (80-85 aa) (Lozano et al. 2017). The RNA-CP

binding motif (RFAAFDFFDAV/KFAAFDFFDAV, 141-151 aa) was detected in the CP sequence (**Fig. 1**) (Alvarez-Quinto et al. 2017).

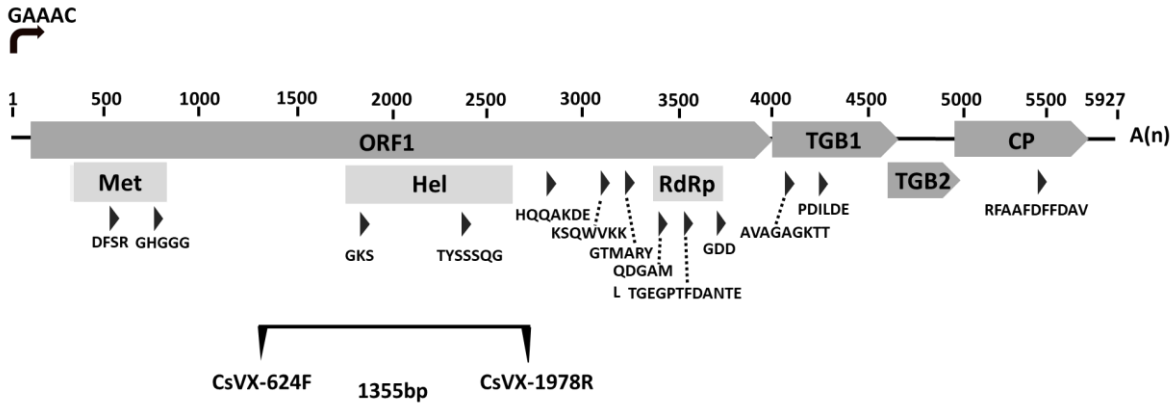
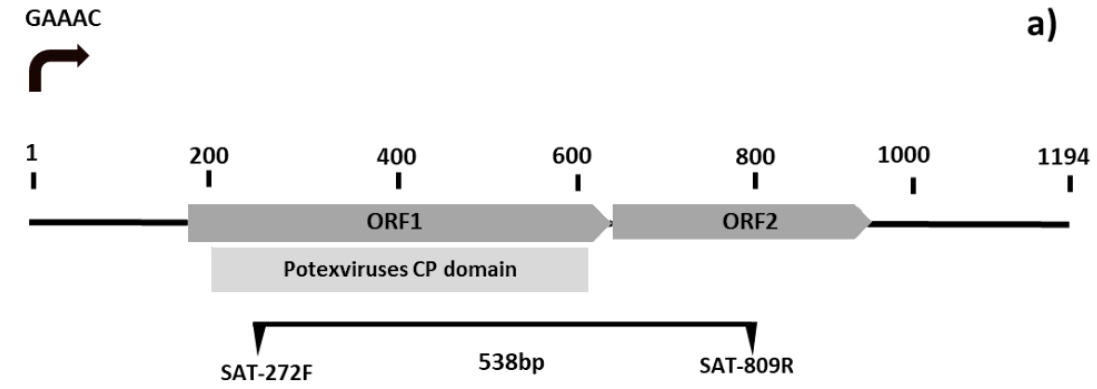
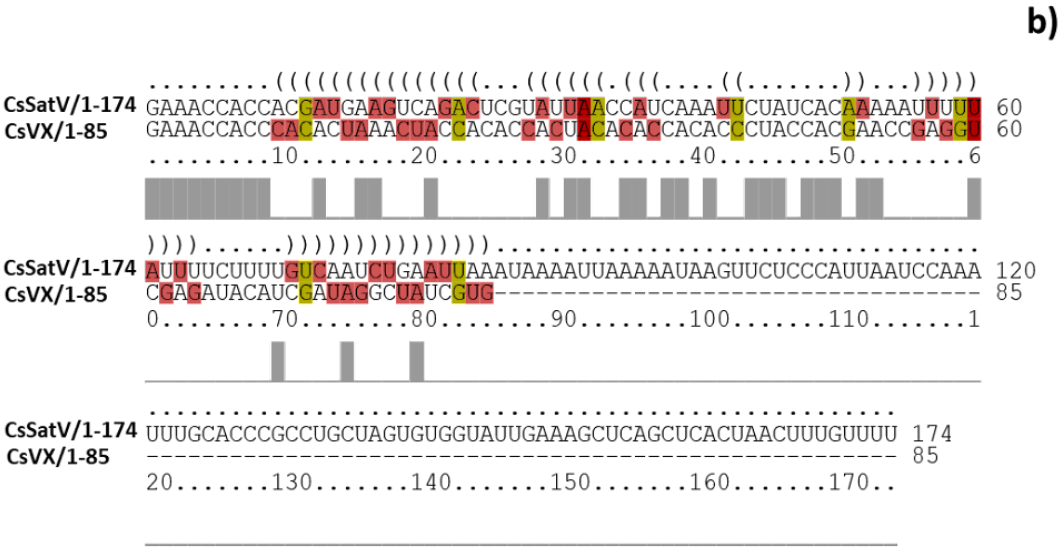


Fig. 1 Schematic representation and characterization of the genome organization of cassava virus X, based on the AJB1_S86 sequence. The genome sequence starts with the typical potexvirus motif (GAAAC). ORFs are represented by a dark gray box indicating the viral sense (+) orientation. Conserved domains of Methyltransferase (Met), Helicase (Hel) and RNA-dependent RNA polymerase (RdRp) with motifs marked with black triangles. Primers (CsVX-624F and CsVX-1987R) used to amplify a 1355 bp fragment for Sanger sequencing are shown.

From the RT-PCR pooled samples, only the one containing samples CPATU 147, CPATU 448, CPATU 452, CPATU 390, CPATU 390was positive for CsVX. Further investigation demonstrated that only samples from Bahia Agará were positive for CsVX and the cassava satellite virus. The CsVX amplified fragment of 1355 bp and the cassava satellite fragment (538 bp) were confirmed by Sanger sequencing. The cassava satellite virus genomic sequence is 1204 nt and has two ORFs. ORF1 is 471 nt long and putatively codes for a protein of 156 aa. An overlapping AUGA start/stop codons were found at the end of ORF1. A conserved domain of potexvirus CP was detected in the ORF1 aa sequence with SMART (Pfam database, **Fig. 2**). The ORF2 is 273 nt and putatively codes for a protein of 90 aa (**Fig. 2**). Blastn with discontiguous megablast mode using the complete assembled cassava satellite virus sequence resulted in 95.08% nt identity (query cover 100%) with the Brazilian sequence of cassava satellite (KY607769) and 73.38% (cover 29%), respectively, with the grapevine satellite virus sequence from USA (MH802033). Similar results were obtained with Blastp searchers. The

ORF1 aa sequence revealed 97.08% (cover of 87%) identity with the ORF1 aa sequence of the Brazilian cassava satellite virus (YP_009345894), followed by 62.22% (cover of 85%) with the grapevine satellite virus (QBZ78379) ORF1 aa sequence from USA. The ORF2 aa sequence shares 100% identity with the ORF2 of the cassava satellite virus followed by 38.46% identity with the grapevine satellite virus sequence (YP_008083723). Sequence alignment of the first 53 nt of 5'-UTRs of the Brazilian CsVX and cassava satellite virus (**Fig. 2**) showed 54,7% identical sites, indicating the possibility that CsVX may act as the helper virus. A predicted consensus secondary structure of the aligned 5´-UTR of CsVX and the cassava satellite virus is shown at **Figure 2**.


a)
GAAAC
1
200
400
600
800
1000
1194
ORF1
ORF2
Potexviruses CP domain
SAT-272F
538bp
SAT-809R


b)
CsSatV/1-174
CsVX/1-85
GAAACCACCACGAUGAAGUCAGACUCGUAUUAACCAUCAAAUUCUAUCACAAAAAUUUU 60
GAAACCACCCACACUAAACUACCACACCACUACACACCACACCCUACCACGAACCGAGGU 60
AUUUUCUUUUGUCAAUCUGAAUUAAAUAAAAUUAAAAAUAAGUUCUCCCAUUAAUCCAAA 120
CGAGAUACAUCCAUAGGCUAUCGUG---------- 85
UUUGCACCCGCCUGCUAGUGUGGUAUUGAAAGCUCAGCUCACUAACUUUGUUUU 174
---------- 85

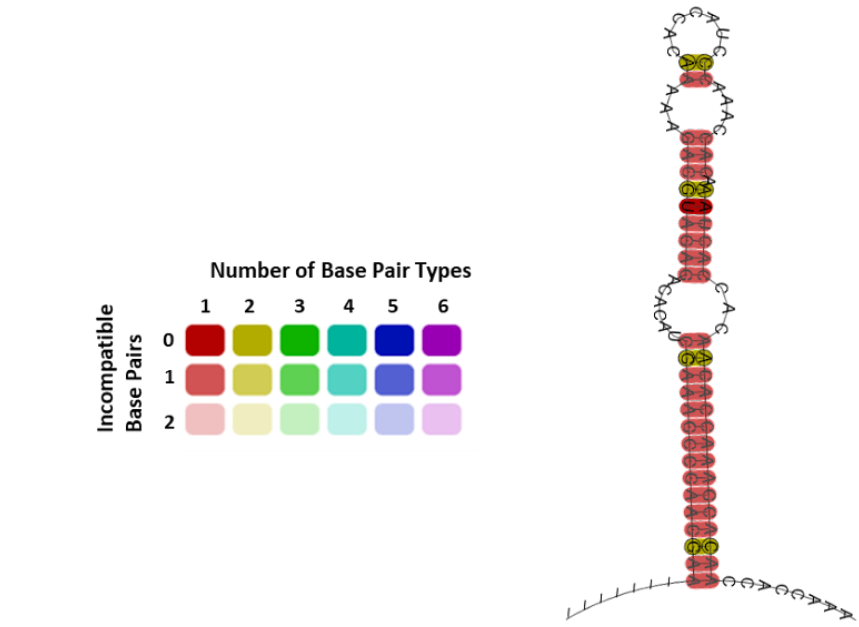

c)
Number of Base Pair Types
1 2 3 4 5 6
Incompatible Base Pairs
0
1
2

Fig. 2 Schematic representation and characterization of the cassava satellite virus sequence found in the AJB1_S86 HTS (A) 5'-UTR structural alignment showing the predicted consensus secondary structure (dot bracket annotation) (B) and conserved structure with mutational pattern. Colors indicate levels of incompatible pairs according to the RNAalifold schema: dark red color indicates high compatibility and conservation according to the legend (C). Genomic sequence starts also with the GAAAC motif, as found for CsVX. ORFs are represented by a dark gray box indicating the satellite sense (+) orientation. Conserved domain of CP is represented with a light gray box. Primer SAT-272F and SAT-809R were used to amplify a 538 bp fragment and are shown below the schematic picture.

SDT sequence comparisons and Bayesian inference

SDT comparisons of all CsVX sequences showed that nt and aa identities varied from 82.9-89.6% and 94.6-98.9%, respectively, in the ORF1 region, 75.1-89.6% and 80.4-94.2% in the TGB1 and 69.0-82.3% and 73.7-91.9% in the TGB2 region, respectively. Comparisons performed with the CP region showed that CsVX isolates shared nt and aa identities that ranges from 81.3-91.8% and 93.0-97.7%, respectively (**Supplementary file 1**). The Colombian CsVX sequence (SRR1050897) shares 87.2 nt and 97.2% aa identities with the sequence from Venezuela (NC_034375) in the ORF1 region, and 84.7 nt and 96.7% aa, respectively, in the CP region. According to the SDT analyses (**Table 2**), the Brazilian sequence (AJB1_S86) showed a lower percentage of identities with the Venezuelan sequence, of 82.9 (nt) and 94.9% (aa) in the ORF1 and 82.3 (nt) and 97.2% (aa) in the CP regions.

Table 2 Sequence lengths and pairwise sequence identity scores from SDT analysis (nt and aa) of the ORF1, TGB1, TGB2 and CP regions of cassava potexviruses. Comparisons were performed using the cassava virus X sequence from Venezuela (NC_034375) as reference.

length (nt)	Sequence	Isolate	ORF1	TBG1	TBG2	CP
	AJB1_S86 cassava virus X	Brazilian	3930	699	300	648
	SRR1050897 cassava virus X	Colombian	3912	693	300	648
	SRR10480878 cassava virus X	Chinese	3915	693	300	648
	NC_034375 cassava virus X	Venezuelan	3915	693	300	645
% identity (nt)	Sequence	Isolate	ORF1	TBG1	TBG2	CP
	KT002439 cassava common mosaic virus	isolate Arg1271	55,0	51,3	50,8	53,7
	MT279196 cassava common mosaic virus	isolate BR	55,5	50,3	56,9	53,9
	NC_001658 cassava common mosaic virus	Brazilian	55,0	55,0	52,0	55,6
	MW175326 cassava common mosaic virus	Hainan-CM	55,8	53,5	55,7	53,8
	MN243731 cassava common mosaic virus	Hainan-DZ	55,5	53,8	52,3	55,6
	MT038420 cassava common mosaic virus	LG	54,8	54,6	58,9	53,5
	KT002435 cassava common mosaic virus	Mcol221	55,0	53,2	51,2	55,8
	MT334616 cassava common mosaic virus	PisVera	57,8	58,0	52,4	42,5
	MN428639 cassava common mosaic virus	China	55,3	53,6	52,9	55,2
	AJB1_S86 cassava virus x	Brazilian	82,9	75,9	74,3	82,3
	SRR1050897 cassava virus x	Colombian	87,2	83,8	79,7	84,7
	SRR10480878 cassava virus x	Chinese	86,9	83,0	82,3	83,6
	KY288505 cassava virus x	CM546010*	88,4	82,2	78,7	83,7
	KC505252 cassava colombian symptomless virus	Colombian	64,8	53,3	60,0	67,4
% identity (aa)	Sequence	Isolate	ORF1	TBG1	TBG2	CP
	KT002439 cassava common mosaic virus	isolate Arg1271	48,8	34,1	22,2	41,1
	MT279196 cassava common mosaic virus	isolate BR	48,5	33,2	25,5	42,5
	NC_001658 cassava common mosaic virus	Brazilian	48,1	32,7	21,2	43,0

MW175326 cassava common mosaic virus	Hainan-CM	48,1	33,2	25,5	43,5
MN243731 cassava common mosaic virus	Hainan-DZ	48,5	32,7	21,2	42,5
MT038420 cassava common mosaic virus	LG	48,4	32,7	25,5	41,6
KT002435 cassava common mosaic virus	Mcol221	48,3	32,7	23,7	43,0
MT334616 cassava common mosaic virus	PisVera	52,4	32,3	22,2	42,5
MN428639 cassava common mosaic virus	China	48,5	32,7	21,2	43,0
AJB1_S86 cassava virus x	Brazilian	94,9	81,3	78,8	97,2
SRR1050897 cassava virus x	Colombian	97,2	91,3	83,8	96,7
SRR10480878 cassava virus x	Chinese	97,3	88,7	87,9	96,7
KY288505 cassava virus x	CM546010	98,9	88,8	80,8	94,4
KC505252 cassava colombian symptomless virus	Colombian	71,3	50,0	55,7	77,1

Bayesian inference was performed using the ORF1 aa and CP coding regions of CsVX. Runs were stopped with an average standard deviation of the division frequencies of 0.000864 and 0.003290, respectively. Phylogenetic analysis showed the formation of a single clade including all CsVX sequences and the cassava Colombian symptomless virus sequence, suggesting a close phylogenetic relationship (**Fig. 3 and Fig. 4**). The tanglegram of trees revealed topological concordance between CsVX and CsCSV clades. CsVX shows a closer relationship to CsCSV than to CsCMV (**Supplementary file 2**).

Fig. 3 Bayesian phylogenetic reconstruction of potexviruses based on the amino acid alignment of the ORF1 (1294 aa aligned sites).

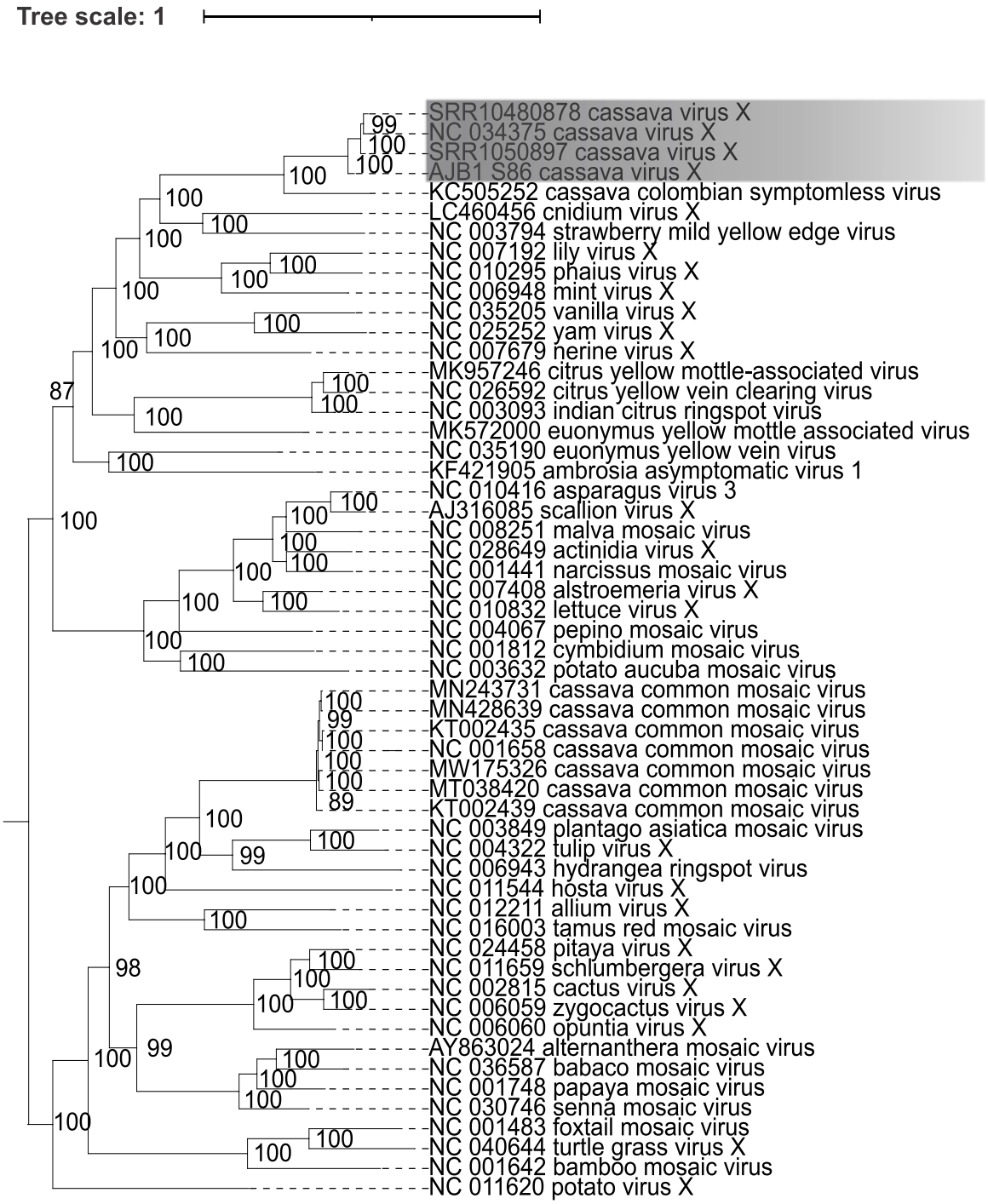
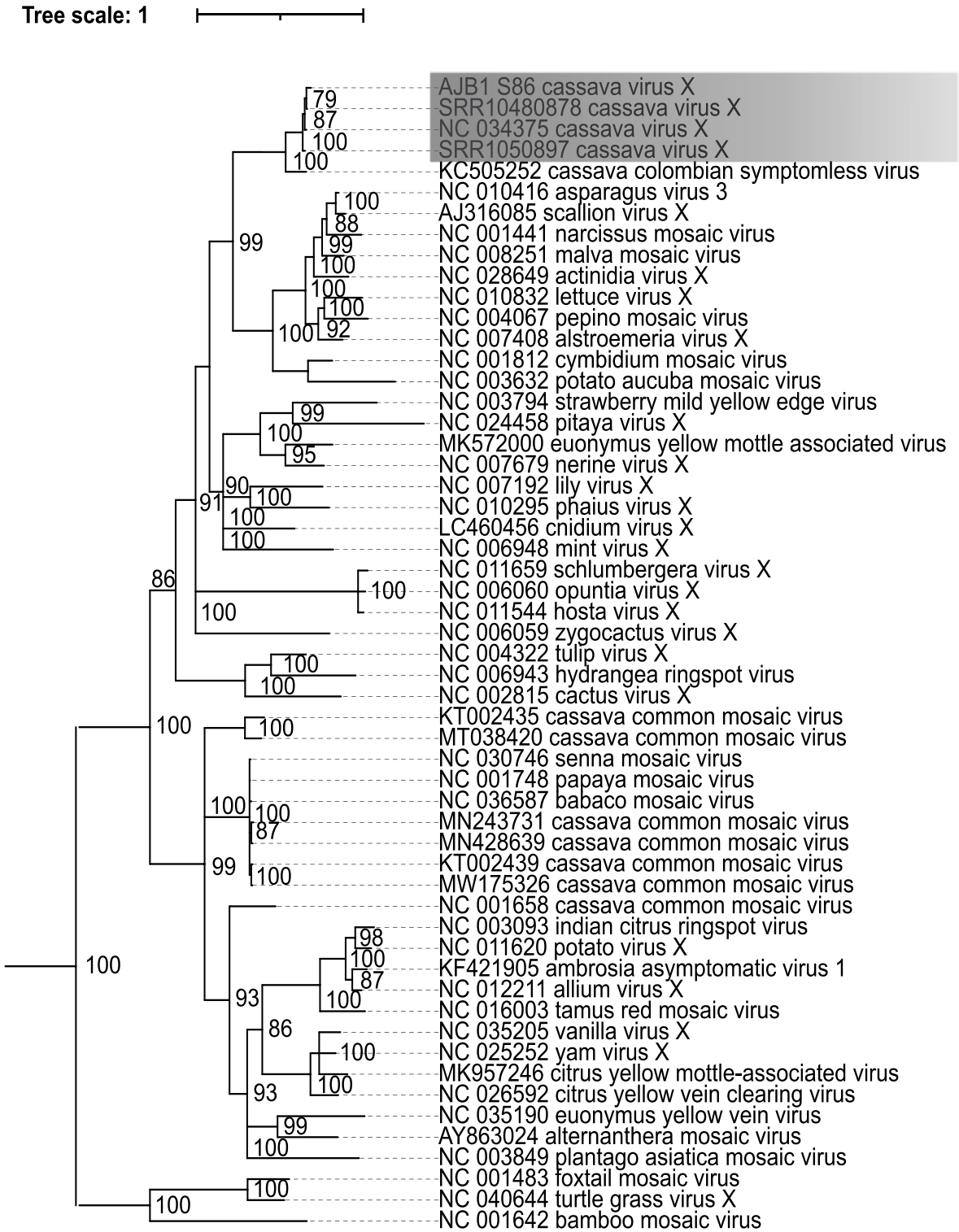


Fig. 4 Bayesian phylogenetic reconstruction of potexviruses based on the amino acid alignment of the CP (205 aa aligned sites).



Discussion

The *Potexvirus* genus is composed of 49 species approved by the ICTV, of which three species have been reported infecting infect cassava plants: cassava common mosaic virus, cassava colombian symptomless virus and cassava virus X (Kreuze et al. 2020). CsCSV and CsVX are known to lack the TGB3 gene (Lozano et al. 2017). CsVX was first reported infecting cassava in 1985 (Harrison et al. 1986, Lennon et al. 1985). In 1992, CsVX was first described in mixed infections in South America (Colombia) infecting cassava plants expressing symptoms of Frog Skin Disease (FSD) (Carvajal-Yepes et al. 2014), and earlier reports show that CsVX transmitted by mechanical inoculation from herbaceous test plants to cassava cultivars quickly develop symptoms in leaves, but not in directly inoculated cassava plants (Lennon et al., 1985). Similar results were shown by Chaparro-Martínez and Trujillo-Pinto in 2001, in which cultivated cassava plants did not express symptoms in the field, although the transmission of CsVX through infected cuttings was 100% effective. Also, Lozano et al. (2017) reported that cassava plants infected with CsVX and CsCSV did not express symptoms in mechanical transmission experiments. According to the authors the only potexvirus that causes disease in single infections is CsCMV, which has been found in cassava plants in Brazil (Costa and Kitajima, 1972; Kitajima et al. 1965) and *Cnidoscolus chayamansa* in Tuvalu and Venezuela (Jones et al. 1998; Mejías et al. 2015). Symptoms of CsVX are, however, expressed in mechanically inoculated *Chenopodium quinoa* and with lower efficiency in *Nicotiana benthamiana* plants (Lennon et al. 1986, Lozano et al. 2017). Therefore, it is still necessary to understand the dynamics of CsVX and CsCSV in mixed infections in cassava and their relationship to the host.

In this work we report for the first time CsVX in cassava plants in the Brazilian Amazon region. We also used SRA data to assemble near-complete genome sequences from Colombia and China. Since isolates of the same species of the genus *Potexvirus* have more than 72% nt or 80% aa identity when compared to their CP or RdRp genes (King et al. 2012), the three assembled sequences are classified as isolates of CsVX. The Brazilian sequence showed a lower nt identity to the sequence from Venezuela than from China and Colombia. Therefore, sequence identity comparisons showed that diversity is not related to geographical location, since the Chinese isolate of CsVX is more closely related to other South American isolates than the Brazilian isolate. Bayesian analysis showed the formation of a single clade of CsVX sequences

with CsCSV, and therefore, they are phylogenetically closer than the CsCMV isolates. These results are in agreement with previous studies (Lozano et al. 2017; Kreuze et al. 2020).

In this work a cassava satellite virus sequence was assembled from the same plant in which CsVX was found. The satellite sequence contains two ORFs. The ORF1 aa sequence has a potexvirus CP domain that is also found in sequences of bamboo mosaic virus satellite RNA (BaMV, Q6XXR6) and panicum mosaic satellite virus (1STM_E). Bamboo mosaic satellite RNA (satBaMV) depends on the bamboo mosaic potexvirus for its replication and encapsidation. However, satBaMV RNA can undergo autonomous long-distance movement in the absence of BaMV (Chang et al. 2016), and it contains a single ORF encoding a protein that has an arginine-rich motif at its N-terminal and shares little homology to BaMV, except for its 5'-UTR. Satellite BaMV mimics the secondary structures of 5′- and 3′-UTRs of BaMV, in which the conserved apical hairpin stem loop (AHSL) downregulates BaMV replication (Lin and Lin, 2017). Interestingly, we also found relatively high sequence identity between the first 53 nt of 5´-UTRs of CsVX and cassava satellite, and a consensus secondary structure from the 5´-UTRs. Therefore, it is tempting to suggest that CsVX may act as the helper virus for cassava satellite virus. CsVX lacks the TGB3 gene and seems to be dispensable for systemic infection of cassava (Lozano et al. 2017), having rather an important effect in cell-to-cell movement than in systemic movement. Future biological experiments should be performed to confirm if cassava satellite virus plays a role in long-distance movement in CsVX infected cassava plants, symptomatology, or even viral replication.

Last, cassava-infecting potexviruses have re-emerged in recent years in South America (Collavino et al. 2021; Zanini et al. 2018) and China (Tuo et al. 2020; You et al. 2020). CsVX was found in HTS data from China showing high nt identities to sequences from South America. Despite the importance of exchanging cassava germplasm materials, it is important to avoid new viral introductions, especially from those latent where symptom expression might depend on specific environmental conditions. It is possible, for instance, that latent infections caused by potexviruses and/or torradoviruses might trigger root symptoms in cassava under rainforest tropical conditions in South America. Finally, this work strives to shed light on potexvirus diversity in cassava in the Amazonian tropical environment.

Author contribution

Conceptualization: AJB, RB, SW; Methodology: GRA, AJB, RB; Formal analysis and investigation: GRA, JACMS, AJB, RB; Writing - original draft preparation: GRA, AJB, RB, SW; Writing - review and editing: GRA, JACMS, AJB, RB, SW; Funding acquisition: AJB, RB, SW; Resources: AJB, RB; Supervision: RB.

Data Availability Statement

The datasets generated and analyzed during the current study are available from the corresponding author.

Acknowledgments

We thank the CAPES (Coordination for the Improvement of Higher Education Personnel) for the scholarship granted to the first author. We also thank the Crop Clinic of Pernambuco from the Rural Federal University of Pernambuco for supporting this work.

Funding

This work was supported by Rural Federal University of Pernambuco (Edital 015/2018-UFRPE PRPPG).

References

Alabi OJ, Kumar PL, Naidu RA (2011) Cassava Mosaic Disease: A Curse to Food Security in Sub-Saharan Africa. APSnet Feature. Available at: https://www.apsnet.org/edcenter/apsnetfeatures/Pages/cassava.aspx#:~:text=Among%20them%2C%20cassava%20mosaic%20disease,of%20leaves%20and%20plants%20. Accessed on February 15, 2022

Alvarez-Quinto RA, Cornejo-Franco JF, Quito-Avila DF (2017) Characterization of a not so new potexvirus from babaco (Vasconcellea x heilbornii). PloS one 12:e0189519

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko I, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev

MA, Pevzner PA (2012) SPAdes: A New Genome Assembly Algo-rithm and Its Applications to Single-Cell Sequencing. Journal of Computational Biology 19: 455-477

Blawid R, Silva JMF, Nagata T (2017) Discovering and sequencing new plant viral genomes by next-generation sequencing: description of a practical pipeline. Annals of Applied Biology 170: 301-314

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30: 2114-2120

Carvajal-Yepes M, Olaya C, Lozano I, Cuervo M, Castaño M, Cuellar WJ (2014) Unraveling complex viral infections in cassava (*Manihot esculenta* Crantz) from Colombia. Virus Research 186:76–86

Chang CH, Hsu FC, Lee SC, Lo YS, Wang JD, Shaw J, Taliansky M, Chang BY, Hsu YH, Lin NS (2016) The Nucleolar Fibrillarin Protein Is Required for Helper Virus-Independent Long-Distance Trafficking of a Subviral Satellite RNA in Plants. The Plant Cell 28:2586-2602.

Chaparro-Martínez EI, Trujillo-Pinto G (2001) First Report of Cassava Virus X in Cassava in Venezuela. Plant Disease 85:1119

Collavino A, Zanini AA, Medina R, Schaller S, Di Feo LDV (2021) Cassava common mosaic virus infection affects growth and yield components of cassava plants (Manihot esculenta) in Argentina Plant Pathology 71: 980-989

Costa AS (1940) Observações sobre o mosaico comum e o mosaico das nervuras da mandioca (*Manihot utilissima* Pohl.). Jornal de Agronomia, 3: 239-248

Darriba D, Taboada GL, Doallo R, Posada D (2011) ProtTest 3: fast selection of best-fit models of protein evolution. Bioin-formatics 8:1164-1165

Dolja VV, Boyko VP, Agranovsky AA, Koonin EV (1991) Phylogeny of capsid proteins of rod-shape/0pçd and filamentous RNA plant viruses: two families with distinct patterns of sequence and probably structure conservation. Virology 184:79-86

Doyle JJ, Doyle JL (1990). Isolation of plant DNA from fresh tissue. Focus, 12: 13-15

FAO. Food and Agriculture Organization of the United Nations. Production of Cassava: top 10 producers 2020. Roma: Food and Agriculture Organization of the United Nations. Available at: http://www.fao.org/faostat/en/#data/QC/visualize. Accessed on January: 10 jan. 2022.

Fedorkin ON, Solovyev AG, Yelina NE, Zamyatnin AA, Zinovkin RA, Mäkinen K, Schiemann J, Yu Morozov S (2001) Cell-to-cell movement of potato virus X involves distinct functions of the coat protein. The Journal of general virology 82:449-458

Harrison BD, Lennon AM, Aiton MM (1986) Properties of four previously uncharacterized cassava viruses. Phytopathology 76:1075

Huson DH, Scornavacca C (2012) Dendroscope 3: An interactive tool for rooted phylogenetic trees and networks. Systematic Biology 61:1061-1067

ICTV. Virus Taxonomy: The Classification and Nomenclature of Viruses The Online (10th), report of the ICTV. Available at: https://talk.ictvonline.org/ictv-reports/ictv_online_report/positive-sense-rna-viruses/w/alphaflexiviridae/1330/genus-potexvirus. Accessed on February 15, 2022

Jones P, Devonshire J, Dabek A, Howells C (1998) First report of Cassava common mosaic potexvirus infecting Chaya (*Cnidoscolus chayamansa*) in Tuvalu. Plant Disease 82: 591

Katoh K, Rozewicki J, Yamada KD (2019) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform 20:1160-1166

King AMQ, Adams MJ, Carstens EB, Lefkowitz E (2012) Virus taxonomy: classification and nomenclature of viruses. Ninth report of the International Committee on Taxonomy of Viruses, Vol. 9. Elsevier, Amsterdam

Kitajima EW, Wetter C, Oliveira AR, Silva DM, Costa AS (1965) Morfologia do vírus do mosaico comum da mandioca. Bragantia 24:247-260

Kreuze JF, Vaira AM, Menzel W, Candresse T, Zavriev SK, Hammond J, Hyun Ryu K, Report Consortium I (2020) ICTV Virus Taxonomy Profile: Alphaflexiviridae. The Journal of general virology, 101:699-700

Lennon AM, Aiton MM, Harrison BD (1986) Cassava viruses from South America. Annual Report 1985, Scottish Crop Research Institute 1985:167

Letunic I, Khedkar S, Bork P (2021) SMART: recent updates, new developments and status in 2020. Nucleic Acids Research 49:D458–D460

Li YI, Shih TW, Hsu YH, Han YT, Huang YL, Meng M (2001) The helicase-like domain of plant potexvirus replicase participates in formation of RNA 5' cap structure by exhibiting RNA 5'-triphosphatase activity. Journal of Virology 75:12114-12120

Lin KY, Lin NS (2017) Interfering Satellite RNAs of Bamboo mosaic virus. Frontiers in Microbiology 4:1-9.

Lough TJ, Emerson SJ, Lucas WJ, Forster RL (2001) Trans-complementation of long-distance movement of White clover mosaic virus triple gene block (TGB) mutants: phloem-associated movement of TGBp1. Virology 288:18-28.

Lozano I, Leiva AM, Jimenez J, Fernandez E, Carvajal-Yepes M, Cuervo M, Cuellar WJ. (2017) Resolution of cassava-infecting alphaflexiviruses: Molecular and biological characterization of a novel group of potexviruses lacking the TGB3 gene. Virus Research 15:53-61

Mejías, A, Rodríguez-Román E, Romano M, Zambrano K, Marys E (2015) New Record of Cassava common mosaic virus Infecting Chaya (*Cnidoscolus chayamansa* McVaug) in Venezuela. Plant Disease 99:1190

McCallum EJ, Anjanappa RB, Gruissem W (2017) Tackling agriculturally relevant diseases in the staple crop cassava (*Manihot esculenta*). Current Opinion in Plant Biology 38:50-58

Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, Tosatto SCE, Paladin L, RAJ S, Richardson LJ, Finn RD, Bateman A (2021) Pfam: The protein families database in 2021. Nucleic Acids Research 49:D412-D419

Morozov SY, Solovyev AG (2003) Triple gene block: modular design of a multifunctional machine for plant virus movement. Journal of General Virology 84:1351-1366

Muhire BM, Varsani A, Martin DP (2014) SDT: A Virus Classification Tool Based on Pairwise Sequence Alignment and Identity Calculation. PLoS ONE 9:e108277

Nolt B, Pineda B, Velasco AC (1992) Surveys of cassava plantations in Colombia for virus and virus-like diseases. Plant Pathol 41:348–354

Oliveira SAS, Ferreira CF, Diamantino MSAS, Santos TA, Pereira JD, Oliveira EJ (2020) First report of cassava torrado-like virus, cassava polero-like virus and cassava new alphaflexivirus associated with cassava frogskin disease in Brazil. Journal of Plant Pathology. 10: 247

Park MR, Jeong RD, Kim KH (2014) Understanding the intracellular trafficking and intercellular transport of potexviruses in their host plants. Frontiers in plant science 5:60

Pecman A, Kutnjak D, Gutiérrez-Aguirre I, Adams I, Fox A, Boonham N, Ravnikar M (2017) Next generation sequencing for detection and discovery of plant viruses and viroids: comparison of two approaches. Frontiers in Microbiology 8:1998

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna 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. Systematic biology, 61: 539–542

Schwarz G (1978) Estimating the Dimension of a Model. Annals of Statistics 6:461-464

Tuo DC, Zhao GY, Yan P, Li RM, Chen X, Wang WQ, Li XY, Yao Y, Cui HG, Shen WT, Zhou P (2020) First Report of Cassava common mosaic virus Infecting Cassava in Mainland China. Disease notes 104:3

UniProt: the universal protein knowledgebase in 2021 (2021) Nucleic Acids Research 49:D480-D489

van der Vlugt, RA, Berendsen M (2002) Development of a General Potexvirus Detection Method. European Journal of Plant Pathology 108:367-371

Venturini MT, Araújo TS, Abreu EFM, Andrade EC, Santos VS, Silva MR, Oliveira EJ (2016) Crop losses in Brazilian cassava varieties induced by the Cassava common mosaic virus. Scientia Agricola 73:520–524

Verchot-Lubicz J (2005) A new cell-to-cell transport model for Potexviruses. Molecular Plant-Microbe Interactions18:283-290

Yu N, Yang Y, Li J, Li W, Liu Z (2020) Complete Genome Sequence of a Distinct Isolate of Cassava common mosaic virus (CsCMV) Infecting Cassava in Hainan, China. Preprints 2020:2020030336

Zanini AA, Cuellar WJ, Celli MG, Luque AV, Medina RD, Conci VC, del Di Feo LV (2018) Distinct strains of the re-emergent cassava common mosaic virus (genus: *Potexvirus*) infecting cassava in Argentina. Plant Pathology 67:1814-1820

Zinga I, Chiroleu F, Legg J, Lefeuvre P, Komba EK, Semballa S, Yandia SP, Mandakombo NB, Reynaud B, Lett J (2013) Epidemiological assessment of cassava mosaic disease in Central African Republic reveals the importance of mixed viral infection and poor health of plant cuttings. Crop Protection 44:6-12

Supplementary file 1 Distance matrix of SDT analysis of the ORF1, TGB1, TGB2 and CP (nt and aa) of cassava potexviruses.

ORF1 nt	KC505252 CsCSV	KT002439 CsCMV	MT279196 CsCMV	NC_001658 CsCMV	MW175326 CsCMV	MN243731 CsCMV	MT038420 CsCMV	KT002435 CsCMV	MT334616 CsCMV	MN428639 CsCMV	AJB1_S86 CsVX	SRR1050897 CsVX	SRR10480878 CsVX	KY288505 CsVX	NC_034375 CsVX
KC505252 cassava colombian symptomless virus	100,0														
KT002439 cassava common mosaic virus (isolate Arg1271)	55,1	100,0													
MT279196 cassava common mosaic virus (isolate BR)	56,0	89,4	100,0												
NC_001658 cassava common mosaic virus (isolate Brazilian)	56,6	88,5	88,5	100,0											
MW175326 cassava common mosaic virus (isolate Hainan-CM)	55,4	89,4	94,7	88,2	100,0										
MN243731 cassava common mosaic virus (isolate Hainan-DZ)	55,5	88,7	88,7	90,1	88,2	100,0									
MT038420 cassava common mosaic virus (isolate LG)	55,6	89,2	93,3	88,2	93,7	88,6	100,0								
KT002435 cassava common mosaic virus (isolate Mcol221)	56,0	88,9	88,8	98,5	88,5	90,4	88,8	100,0							
MT334616 cassava common mosaic virus (isolate PisVera)	56,6	88,9	88,4	93,7	88,3	91,0	88,7	94,1	100,0						
MN428639 cassava common mosaic virus (isolate China)	55,2	88,7	88,7	90,4	88,2	98,9	88,6	90,7	91,0	100,0					
AJB1_S86 cassava virus x (Brazil)	65,8	54,6	54,8	55,0	54,9	54,3	53,8	55,1	56,8	54,3	100,0				
SRR1050897 cassava virus x (Colombia)	66,6	55,2	54,9	55,7	55,3	54,3	55,0	55,2	56,9	54,5	83,5	100,0			
SRR10480878 cassava virus x (China)	64,7	55,9	54,7	55,4	55,0	55,2	54,2	55,0	56,4	55,2	83,0	87,4	100,0		
KY288505 cassava virus x (isolate CM546010*)	67,0	61,5	60,1	51,3	60,6	60,4	59,3	60,0	60,1	60,4	83,1	89,6	86,7	100,0	
NC_034375 cassava virus x (isolate Ven164)	64,8	55,0	55,5	55,0	55,8	55,5	54,8	55,0	57,8	55,3	82,9	87,2	86,9	88,4	100,0

ORF1 aa	KC505252 CsCSV	KT002439 CsCMV	MT279196 CsCMV	NC_001658 CsCMV	MW175326 CsCMV	MN243731 CsCMV	MT038420 CsCMV	KT002435 CsCMV	MT334616 CsCMV	MN428639 CsCMV	AJB1_S86 CsVX	SRR1050897 CsVX	SRR10480878 CsVX	KY288505 CsVX	NC_034375 CsVX
KC505252 cassava colombian symptomless virus	100,0														
KT002439 cassava common mosaic virus (isolate Arg1271)	47,9	100,0													
MT279196 cassava common mosaic virus (isolate BR)	48,1	96,9	100,0												
NC_001658 cassava common mosaic virus (isolate Brazilian)	47,2	95,7	96,4	100,0											
MW175326 cassava common mosaic virus (isolate Hainan-CM)	47,2	97,4	99,1	96,5	100,0										
MN243731 cassava common mosaic virus (isolate Hainan-DZ)	47,8	97,2	98,1	96,6	98,1	100,0									
MT038420 cassava common mosaic virus (isolate LG)	48,2	96,5	97,7	95,2	97,9	97,0	100,0								
KT002435 cassava common mosaic virus (isolate Mcol221)	47,8	96,8	97,5	98,1	97,6	97,8	96,3	100,0							
MT334616 cassava common mosaic virus (isolate PisVera)	50,5	98,7	99,4	98,1	99,6	99,2	98,9	99,1	100,0						
MN428639 cassava common mosaic virus (isolate China)	47,7	97,5	98,3	96,9	98,4	99,7	97,3	98,1	99,4	100,0					
AJB1_S86 cassava virus x (Brazil)	72,1	48,7	48,6	48,3	48,6	48,6	48,4	48,6	52,6	48,6	100,0				
SRR1050897 cassava virus x (Colombia)	72,4	48,7	48,6	48,3	48,6	48,6	48,4	48,5	52,5	48,6	94,9	100,0			
SRR10480878 cassava virus x (China)	72,0	48,8	48,5	48,0	48,5	48,7	48,3	48,3	52,5	48,7	94,6	97,3	100,0		
KY288505 cassava virus x (isolate CM546010*)	73,9	58,3	57,5	57,3	57,5	57,5	57,2	57,7	57,5	57,5	97,0	98,7	98,7	100,0	
NC_034375 cassava virus x (isolate Ven164)	71,3	48,8	48,5	48,1	48,1	48,5	48,4	48,3	52,4	48,5	94,9	97,2	97,3	98,9	100,0

TGB1 nt	KC505252 CsCSV	KT002439 CsCMV	MT279196 CsCMV	NC_001658 CsCMV	MW175326 CsCMV	MN243731 CsCMV	MT038420 CsCMV	KT002435 CsCMV	MT334616 CsCMV	MN428639 CsCMV	AJB1_S86 CsVX	SRR1050897 CsVX	SRR10480878 CsVX	KY288505 CsVX	NC_034375 CsVX
KC505252 cassava colombian symptomless virus	100,0														
KT002439 cassava common mosaic virus (isolate Arg1271)	53,9	100,0													
MT279196 cassava common mosaic virus (isolate BR)	51,1	87,4	100,0												
NC_001658 cassava common mosaic virus (isolate Brazilian)	52,5	86,9	86,9	100,0											
MW175326 cassava common mosaic virus (isolate Hainan-CM)	51,9	87,5	93,0	86,4	100,0										
MN243731 cassava common mosaic virus (isolate Hainan-DZ)	51,8	87,2	87,1	90,8	86,9	100,0									
MT038420 cassava common mosaic virus (isolate LG)	52,4	88,9	93,1	87,4	93,4	87,1	100,0								
KT002435 cassava common mosaic virus (isolate Mcol221)	51,4	87,2	86,9	98,9	86,6	90,7	87,8	100,0							
MT334616 cassava common mosaic virus (isolate PisVera)	51,0	86,9	87,4	93,5	85,8	91,5	86,9	92,7	100,0						
MN428639 cassava common mosaic virus (isolate China)	51,8	87,1	86,9	90,9	86,8	99,9	86,9	90,8	91,4	100,0					
AJB1_S86 cassava virus x (Brazil)	52,0	53,7	51,1	55,4	52,3	52,1	52,3	55,2	53,7	52,0	100,0				
SRR1050897 cassava virus x (Colombia)	55,7	50,9	49,7	54,6	49,7	50,2	50,7	52,5	51,6	50,8	76,6	100,0			
SRR10480878 cassava virus x (China)	54,4	51,3	51,1	50,6	50,2	51,6	51,8	50,8	50,2	51,8	75,5	84,6	100,0		
KY288505 cassava virus x (isolate CM546010)	56,6	50,2	50,3	53,7	52,8	50,8	51,5	52,9	53,7	50,9	75,1	89,6	83,4	100,0	
NC_034375 cassava virus x (isolate Ven164)	53,3	51,3	50,3	55,0	53,5	53,8	54,6	53,2	58,0	53,6	75,9	83,8	83,0	82,2	100,0

TGB1 aa	KC505252 CsCSV	KT002439 CsCMV	MT279196 CsCMV	NC_001658 CsCMV	MW175326 CsCMV	MN243731 CsCMV	MT038420 CsCMV	KT002435 CsCMV	MT334616 CsCMV	MN428639 CsCMV	AJB1_S86 CsVX	SRR1050897 CsVX	SRR10480878 CsVX	KY288505 CsVX	NC_034375 CsVX
KC505252 cassava colombian symptomless virus	100,0														
KT002439 cassava common mosaic virus (isolate Arg1271)	31,5	100,0													
MT279196 cassava common mosaic virus (isolate BR)	30,4	96,5	100,0												
NC_001658 cassava common mosaic virus (isolate Brazilian)	30,8	93,9	95,7	100,0											
MW175326 cassava common mosaic virus (isolate Hainan-CM)	29,9	95,2	98,7	94,4	100,0										
MN243731 cassava common mosaic virus (isolate Hainan-DZ)	29,5	96,1	97,4	96,1	96,1	100,0									
MT038420 cassava common mosaic virus (isolate LG)	30,4	95,2	98,7	94,4	97,4	96,1	100,0								
KT002435 cassava common mosaic virus (isolate Mcol221)	30,8	94,4	96,1	99,6	94,8	96,5	94,8	100,0							
MT334616 cassava common mosaic virus (isolate PisVera)	30,4	94,4	95,2	98,7	93,9	96,5	93,9	99,1	100,0						
MN428639 cassava common mosaic virus (isolate China)	29,5	96,1	97,4	96,1	96,1	100,0	96,1	96,5	96,5	100,0					
AJB1_S86 cassava virus x (Brazil)	49,1	35,9	35,5	35,0	35,5	35,0	35,0	35,0	35,0	35,0	100,0				
SRR1050897 cassava virus x (Colombia)	50,9	32,6	32,1	31,7	32,1	31,7	31,7	31,7	31,2	31,7	82,2	100,0			
SRR10480878 cassava virus x (China)	51,8	33,2	33,6	31,7	33,6	33,2	33,2	31,7	31,2	33,2	82,6	91,3	100,0		
KY288505 cassava virus x (isolate CM546010)	50,9	35,0	34,5	34,1	34,5	34,1	33,6	34,1	33,6	34,1	80,4	94,2	88,8	100,0	
NC_034375 cassava virus x (isolate Ven164)	50,0	34,1	33,2	32,7	33,2	32,7	32,7	32,7	32,3	32,7	81,3	91,3	88,7	88,8	100,0

TGB2 nt	KC505252 CsCSV	KT002439 CsCMV	MT279196 CsCMV	NC_001658 CsCMV	MW175326 CsCMV	MN243731 CsCMV	MT038420 CsCMV	KT002435 CsCMV	MT334616 CsCMV	MN428639 CsCMV	AJB1_S86 CsVX	SRR1050897 CsVX	SRR10480878 CsVX	KY288505 CsVX	NC_034375 CsVX
KC505252 cassava colombian symptomless virus	100,0														
KT002439 cassava common mosaic virus (isolate Arg1271)	48,6	100,0													
MT279196 cassava common mosaic virus (isolate BR)	48,2	95,2	100,0												
NC_001658 cassava common mosaic virus (isolate Brazilian)	47,5	94,9	95,5	100,0											
MW175326 cassava common mosaic virus (isolate Hainan-CM)	47,5	94,9	96,1	95,2	100,0										
MN243731 cassava common mosaic virus (isolate Hainan-DZ)	47,6	93,7	93,7	94,6	92,8	100,0									
MT038420 cassava common mosaic virus (isolate LG)	49,6	93,4	94,6	93,7	94,9	92,5	100,0								
KT002435 cassava common mosaic virus (isolate Mcol221)	47,4	94,3	95,2	99,4	94,9	94,3	93,4	100,0							
MT334616 cassava common mosaic virus (isolate PisVera)	44,6	95,2	95,8	97,3	94,9	95,5	94,0	96,7	100,0						
MN428639 cassava common mosaic virus (isolate China)	47,2	93,1	94,3	94,6	93,4	99,4	93,1	94,3	96,1	100,0					
AJB1_S86 cassava virus x (Brazil)	57,1	51,0	52,2	54,0	53,8	52,6	51,7	54,3	51,8	52,6	100,0				
SRR1050897 cassava virus x (Colombia)	54,9	50,0	51,5	51,6	50,5	53,0	55,7	51,3	51,3	52,2	75,3	100,0			
SRR10480878 cassava virus x (China)	58,7	51,4	55,2	54,2	51,9	53,3	56,1	54,2	55,2	58,6	69,0	78,0	100,0		
KY288505 cassava virus x (isolate CM546010)	64,2	52,8	47,2	51,4	50,2	49,3	52,8	53,8	47,9	47,5	75,7	87,7	78,7	100,0	
NC_034375 cassava virus x (isolate Ven164)	60,0	50,8	56,9	52,0	55,7	52,3	58,9	51,2	52,4	52,9	74,3	79,7	82,3	78,7	100,0

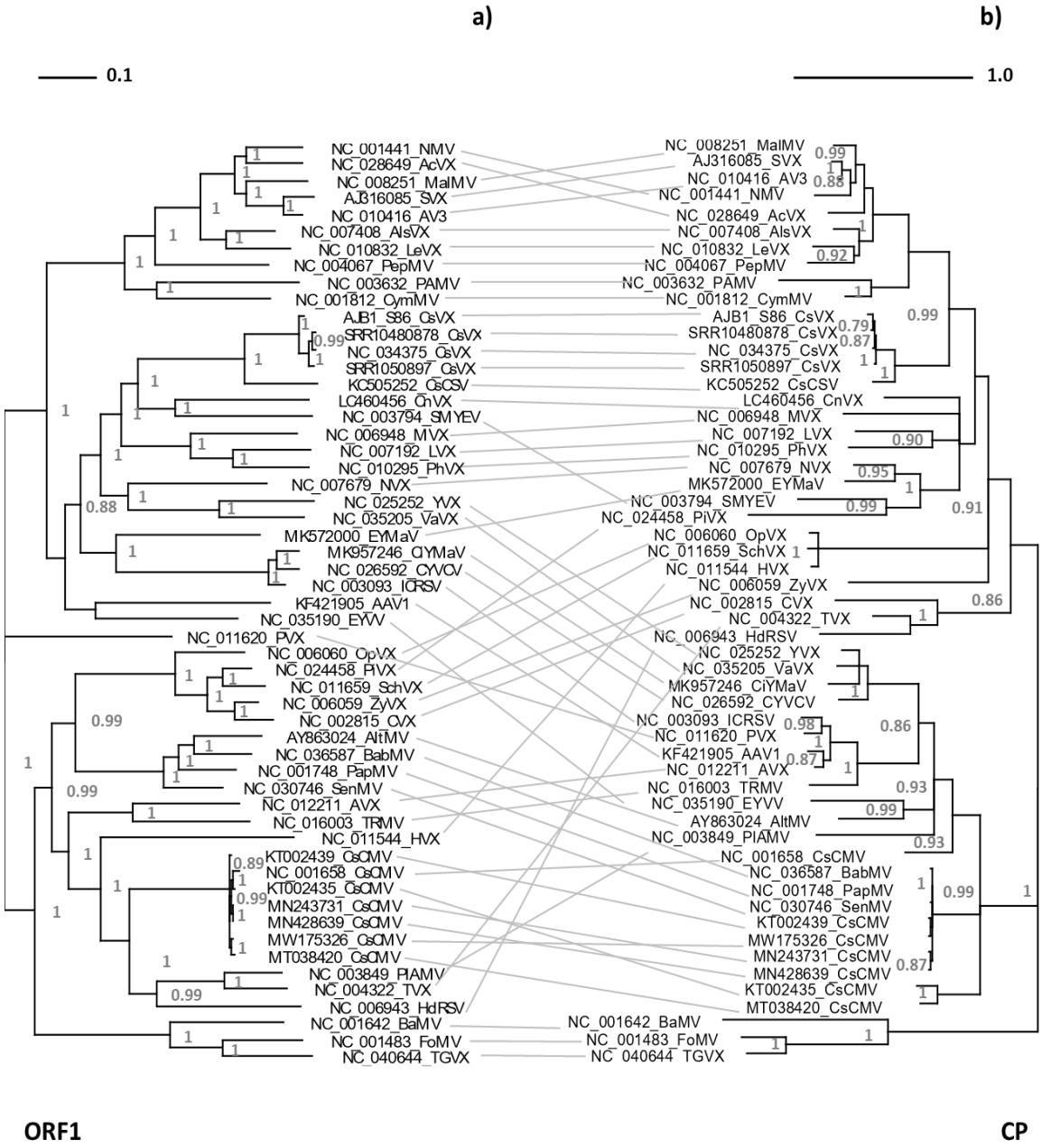
TGB2 aa	KC505252 CsCSV	KT002439 CsCMV	MT279196 CsCMV	NC_001658 CsCMV	MW175326 CsCMV	MN243731 CsCMV	MT038420 CsCMV	KT002435 CsCMV	MT334616 CsCMV	MN428639 CsCMV	AJB1_S86 CsVX	SRR1050897 CsVX	SRR10480878 CsVX	KY288505 CsVX	NC_034375 CsVX
KC505252 cassava colombian symptomless virus	100,001														
KT002439 cassava common mosaic virus (isolate Arg1271)	26,5	100,0													
MT279196 cassava common mosaic virus (isolate BR)	24,5	96,4	100,0												
NC_001658 cassava common mosaic virus (isolate Brazilian)	25,3	97,3	96,4	100,0											
MW175326 cassava common mosaic virus (isolate Hainan-CM)	27,8	95,5	96,4	95,5	100,0										
MN243731 cassava common mosaic virus (isolate Hainan-DZ)	26,5	95,5	95,5	96,4	94,5	100,0									
MT038420 cassava common mosaic virus (isolate LG)	25,5	94,5	96,4	94,5	94,5	93,6	100,0								
KT002435 cassava common mosaic virus (isolate Mcol221)	26,5	95,5	95,5	98,2	94,5	95,5	93,6	100,0							
MT334616 cassava common mosaic virus (isolate PisVera)	25,5	94,5	94,5	96,4	92,7	95,5	93,6	94,5	100,0						
MN428639 cassava common mosaic virus (isolate China)	27,8	94,5	96,4	96,4	94,5	99,1	94,5	95,5	96,4	100,0					
AJB1_S86 cassava virus x (Brazil)	52,0	25,8	24,7	23,2	24,7	24,2	23,7	25,8	23,2	24,2	100,0				
SRR1050897 cassava virus x (Colombia)	54,1	19,6	24,0	20,6	24,0	19,6	24,0	20,6	20,6	19,6	77,8	100,0			
SRR10480878 cassava virus x (China)	53,1	30,2	33,3	33,3	33,3	30,1	33,3	32,3	35,5	32,3	73,7	82,8	100,0		
KY288505 cassava virus x (isolate CM546010)	52,0	22,6	21,5	22,7	18,6	21,5	18,6	21,5	22,6	21,5	78,8	91,9	81,8	100,0	
NC_034375 cassava virus x (isolate Ven164)	55,7	22,2	25,5	21,2	25,5	21,2	25,5	23,7	22,2	21,2	78,8	83,8	87,9	80,8	100,0

CP nt	KC505252 CsCSV	KT002439 CsCMV	MT279196 CsCMV	NC_001658 CsCMV	MW175326 CsCMV	MN243731 CsCMV	MT038420 CsCMV	KT002435 CsCMV	MT334616 CsCMV	MN428639 CsCMV	AJB1_S86 CsVX	SRR1050897 CsVX	SRR10480878 CsVX	KY288505 CsVX	NC_034375 CsVX
KC505252 cassava colombian symptomless virus	100,0														
KT002439 cassava common mosaic virus (isolate Arg1271)	50,5	100,0													
MT279196 cassava common mosaic virus (isolate BR)	53,4	90,6	100,0												
NC_001658 cassava common mosaic virus (isolate Brazilian)	52,7	90,4	89,0	100,0											
MW175326 cassava common mosaic virus (isolate Hainan-CM)	49,8	89,9	93,9	88,8	100,0										
MN243731 cassava common mosaic virus (isolate Hainan-DZ)	49,6	89,7	89,3	92,2	89,1	100,0									
MT038420 cassava common mosaic virus (isolate LG)	50,1	92,2	92,8	88,6	93,3	88,6	100,0								
KT002435 cassava common mosaic virus (isolate Mcol221)	51,9	90,0	89,1	99,3	88,4	91,7	88,1	100,0							
MT334616 cassava common mosaic virus (isolate PisVera)	38,5	71,6	70,3	74,8	70,9	72,0	70,0	74,3	100,0						
MN428639 cassava common mosaic virus (isolate China)	50,1	89,7	89,1	92,3	89,0	98,7	88,4	91,9	72,0	100,0					
AJB1_S86 cassava virus x (Brazil)	67,7	53,4	51,2	52,9	53,7	52,6	52,7	52,4	39,2	52,5	100,0				
SRR1050897 cassava virus x (Colombia)	67,7	52,5	50,5	50,1	52,4	52,2	50,8	50,1	39,0	51,5	81,8	100,0			
SRR10480878 cassava virus x (China)	65,9	55,2	51,1	52,1	53,2	52,7	52,1	52,2	41,8	53,9	81,3	85,5	100,0		
KY288505 cassava virus x (isolate CM546010)	67,1	54,0	52,4	50,5	53,2	51,5	51,0	50,7	39,4	52,6	82,3	91,8	84,7	100,0	
NC_034375 cassava virus x (isolate Ven164)	67,4	53,7	53,9	55,6	53,8	55,6	53,5	55,8	42,5	55,2	82,3	84,7	83,6	83,7	100,0

CP aa	KC505252 CsCSV	KT002439 CsCMV	MT279196 CsCMV	NC_001658 CsCMV	MW175326 CsCMV	MN243731 CsCMV	MT038420 CsCMV	KT002435 CsCMV	MT334616 CsCMV	MN428639 CsCMV	AJB1_S86 CsVX	SRR1050897 CsVX	SRR10480878 CsVX	KY288505 CsVX	NC_034375 CsVX
KC505252 cassava colombian symptomless virus	100,0														
KT002439 cassava common mosaic virus (isolate Arg1271)	38,5	100,0													
MT279196 cassava common mosaic virus (isolate BR)	39,4	97,4	100,0												
NC_001658 cassava common mosaic virus (isolate Brazilian)	38,5	98,3	97,4	100,0											
MW175326 cassava common mosaic virus (isolate Hainan-CM)	39,1	96,9	97,8	98,3	100,0										
MN243731 cassava common mosaic virus (isolate Hainan-DZ)	38,6	96,5	95,2	97,4	96,1	100,0									
MT038420 cassava common mosaic virus (isolate LG)	38,9	97,8	98,7	97,8	98,3	95,6	100,0								
KT002435 cassava common mosaic virus (isolate Mcol221)	39,5	97,4	96,5	99,1	97,4	96,5	96,9	100,0							
MT334616 cassava common mosaic virus (isolate PisVera)	37,6	96,9	96,1	98,7	96,9	96,9	96,5	97,8	100,0						
MN428639 cassava common mosaic virus (isolate China)	39,5	96,5	95,6	97,8	96,5	98,7	96,1	96,9	96,5	100,0					
AJB1_S86 cassava virus x (Brazil)	75,8	41,9	42,8	42,8	42,8	42,3	41,9	42,8	42,3	42,8	100,0				
SRR1050897 cassava virus x (Colombia)	77,2	41,4	42,3	42,3	42,3	41,9	41,9	42,3	41,9	42,3	94,9	100,0			
SRR10480878 cassava virus x (China)	76,3	42,3	43,3	42,3	42,3	41,9	42,3	42,3	41,9	42,3	94,9	95,8	100,0		
KY288505 cassava virus x (isolate CM546010)	76,3	40,9	40,9	41,4	41,4	40,9	40,9	41,4	40,9	41,4	93,0	97,7	94,0	100,0	
NC_034375 cassava virus x (isolate Ven164)	77,1	41,1	42,5	43,0	43,5	42,5	41,6	43,0	42,5	43,0	97,2	96,7	96,7	94,4	100,0

*Partial sequence

Supplementary file 2 A tanglegram depicting phylogenetic potexvirus relationships between inferred Bayesian phylogenetic trees of ORF1 (1294 aa aligned sites) and CP (205 aa aligned sites) regions. Scale bar indicates 0.1 and 1.0 amino acid substitutions per site of ORF1 and CP, respectively.



CAPÍTULO IV

First report of a polerovirus infecting cassava plants from the Amazon region

Archives of Virology-Annotated Sequence Records

First report of a polerovirus infecting cassava plants from the Amazon region.

[1]Géssyka Rodrigues de Albuquerque, [2]Alessandra de Jesus Boari [2]Elisa Ferreira Moura Cunha, [2]Késsia Fátima Cunha Pantoja, [1]Rosana Blawid

1 Universidade Federal Rural de Pernambuco, Recife, Pernambuco, Brazil.
2 Embrapa Amazônia Oriental, Belém, Pará, Brazil.

Abstract
The most economically important disease in Latin America is the Cassava Frog Skin Disease (CFSD) syndrome. Cassava polero-like virus has been reported in mixed infections in cassava plants exhibiting CFSD-like symptoms. In the present work, we characterized a cassava polero-like virus found infecting cassava (*Manihot esculenta* Crantz.) plants of the Amazon region. For this, total RNA was extracted from 14 different cassava varieties kept by the Cassava Germplasm Bank of the Embrapa Eastern Amazon and send for High-Throughput Sequencing. The complete genome was also confirmed by Sanger sequencing from a single plant (variety CPATU 147). After assembling the complete genome, specific primers were designed, and detection was performed in the samples separately. BLASTp analysis of the P1P2 and P3P5 regions showed amino acid identity levels of 95.06% (AHA91815) and 97.35% (ALX37930), respectively, with partial sequences from Colombia. This is the first worldwide report of a complete sequence of a polerovirus infecting cassava plants. This characterization will provide molecular tools for more extensive studies aiming to investigate the putative association of cassava polero-like virus and the CFSD syndrome.

Poleroviruses are plant viruses that infect many economically important crops (Delfosse et al. 2021). The particles of poleroviruses are icosahedral and nonenveloped (Byrne et al. 2019). Genomic positive-sense RNA of poleroviruses is polycistronic, codes up to 10 ORFs, contains 5'UTR, 3'UTR, and an untranslated intergenic region (Delfosse et al. 2021; Jeevalatha et al. 2013; Jaag et al. 2003). In ORF1 translation ribosomes use a leaky scanning mechanism by the ORF0, and a -1 ribosomal frameshift mechanism in ORF1 is used for reading of the ORF2. OFR2 encodes the RNA-dependent RNA polymerase (RdRp). Subgenomic RNA1 is used for the translation of 3a, 3, 4 and 5 ORFs, a subgenomic RNA2 is used for the ORF6 and ORF7 translation, and a subgenomic RNA3 only for the translation of ORF7. ORF3 and ORF5 are

fused and the ORF5 expression occurs by a readthrough mechanism from the stop codon of ORF3 (Delfosse et al. 2021; Sõmera et al. 2021; Smirnova et al. 2015; Jaag et al. 2003). An ORF8 was found in the potato leafroll virus (PLRV), coding only for the replication-associated protein Rap1 (Jeevalatha et al. 2013).

The viral infection in plants is phloem specific and changes in the phloem proteome appear to contribute to viral infection. Poleroviruses are transmission by aphids (Cilia et al. 2012; Gray et al. 2003), and by whitefly (Ghosh et al. 2019). And several symptoms have been associated with viral infection. Symptoms expressed in plants infected by polerovirus include yellowing, reddening, curling, small leaves, discoloration of the main leaf vein, bushy and brittle appearance, internodes shortening and stunting (Abraham et al. 2022; Zhao et al. 2021; Distéfano et al. 2010). Cassava polero-like virus has been reported in mixed infections in cassava plants expressing symptoms of Frogskin Disease (FSD) (Oliveira et al. 2020; Carvajal-Yepes et al. 2014). This study aims to detect and characterize the first complete genome of cassava polerovirus from cassava plant of Embrapa Amazônia Oriental with symptoms of FSD.

Total RNA was isolated from 14 cassava varieties (CPATU 147, CPATU 448, CPATU 452, CPATU 390, CPATU 025, CPATU 016, CPATU 507, CPATU 322, CPATU 232, CPATU 229, CPATU 013, CPATU 250, CPATU 183, and CPATU 592) collected from the Cassava Germplasm Bank area of Embrapa Amazônia Oriental (Amazon, Brazil) using the PureLink$^{@}$ kit for viral RNA/DNA Miniprep (Thermo Fisher Scientific/Invitrogen, MA, USA) following the manufacturer's protocol. The RNA was sequenced (AJB1_S86_R1.fastq and AJB1_S86_R2.fastq) using the Illumina HiSeq 2500 platform with 100-base paired-end reads at the Centro de Genômica Funcional-ESALQ (University of São Paulo, Piracicaba, Brazil). The bioinformatics analyses followed a pipeline for assembly of plant viral genomes described previously by Blawid et al. (2017). Briefly, first the quality of FastQ files of forward and reverse reads were checked with FastQC v. 0.11.9 (http://www.bioinformatics.babraham .ac.uk/projects/fastqc/). Then reads files were trimmed with Trimmomatic v. 0.36. (Bolger et al. 2014) to remove low quality reads and Illumina adapters and used for *de novo* assembled with SPAdes v3.11.1 (Bankevich et al. 2012). tBlastx searchers were performed in Geneious v.11.0.5 using a plant virus database from sequences available at NCBI (9361 sequences, downloaded at 2020). The larger contig (5.938 nt) classified as polerovirus was extracted and extended through consecutive mappings of the reads against the consensus. Genome annotation

of the novel cassava polerovirus was performed in Geneious, with the help of the NCBI BlastP tool (https://blast.ncbi.nlm.nih.gov) and the Simple Modular Architecture Research Tool (SMART) (Letunic et al. 2021). The final assembled cassava polerovirus sequence was used for designing specific primers for sequencing the whole genome by Sanger (Supplementary Table 1). Secondary structures of polerovirus pseudoknot were constructed with IPknot++ web server v.2.1.0 (http://rtips.dna.bio.keio.ac.jp/ipknot++/) and visualized with VARNA v.3 (Darty et al. 2009).

In order to detect the polerovirus on the fourteen cassava genotypes previously mentioned, total RNA was extracted using a modified protocol of the CTAB method (Doyle and Doyle 1990). First, Two-step RT-PCRs were performed from pooled samples (four cultivars) using the Moloney murine leukemia virus (MMLV) reverse transcriptase (Invitrogen™) and oligodeoxyribonucleotides d(N)6 random primers. Standard PCRs were done using the Invitrogen Platinum Taq DNA Polymerase, with the specific primers 2290F and 3343R (Supplementary Table 1). Three pools tested positive and the pooled sample 1 containing four plants was individually inspected. The complete genome was amplified from a single positive plant (variety CPATU 147) with specific primers (Supplementary Table 1). For 5’ and 3'-end amplifications, specific antisense and sense primers, respectively, were used with the SuperScript IV Reverse Transcriptase (Invitrogen™) for cDNA synthesis (Supplementary Table 2), followed by RNaseH (Thermo Scientific™, Waltham, MA) digestion. Tailing reactions were done with addition of dCTP 5mM to cDNA, using the terminal transferase (NEB, MA, USA). To increase the specificity of the reaction, two Nested-PCR amplifications with oligo(dG) and internal specifics primers were performed using the Phusion High-Fidelity PCR Master Mix (Thermo Scientific™, Waltham, MA). Amplicons obtained from the second Nested-PCRs were excised, purified, and sequenced by Sanger sequencing.

For the phylogenetic analysis the following poleroviruses amino acid sequences of P3P5 available at National Center for Biotechnology Information (NCBI) were employed. Multiple sequence alignment was done using MAFFT v.7 (Katoh et al. 2019) implemented in Geneious v11. The Bayesian Information Criterion (BIC) (Schwarz, 1978) was used to estimate the best model of amino acid evolution with ProtTest v.3.4.2 (Darriba et al. 2011). The best model of amino acids was GT+G for the P3-P5 region. Maximum likelihood trees were constructed using PhyML (Guindon et al. 2010) with the P3-P5 aa sequences. Phylogenetic trees were visualized

in the web tool iTOL v6 (Letunic and Bork, 2021). Comparisons between polerovirus sequences were made with the Sequence Demarcation Tool v1.2 (SDT) (Muhire et al. 2014) using MAFFT v.7 alignment.

The dataset AJB1_S86 had a total of 17254734 reads, remaining 15147828 after trimming. The total number of viral contigs generated was 56524, of which 62 were assigned as cassava polerovirus, according to the tBlastX results. The number of reads assembled to the cassava polerovirus was 9971. The genome size of the assembled polerovirus sequence was of 5939 nt (Fig.1).

The P0 protein contains a distinct F-box domain FPLSL (68-72 aa), different from that described for poleroviruses LPXXL/IX(10-13)P (Desfosse et al. 2021). ORF1 (P1-P2 protein) contains the heptanucleotide GGGAAAC sequence, which marks the putative beginning of the pseudoknot of poleroviruses (Desfosse et al. 2021). Additionally, the CP aa sequence motif GVLNAFHE (104-111), and the P5 protein translation starts after the Stop codon (ribosomal read-through) in the sequence 5'-AAAAUAGGUAGAC-3', conserved in poleroviruses were found (Costa et al. 2019) (Fig. 1).

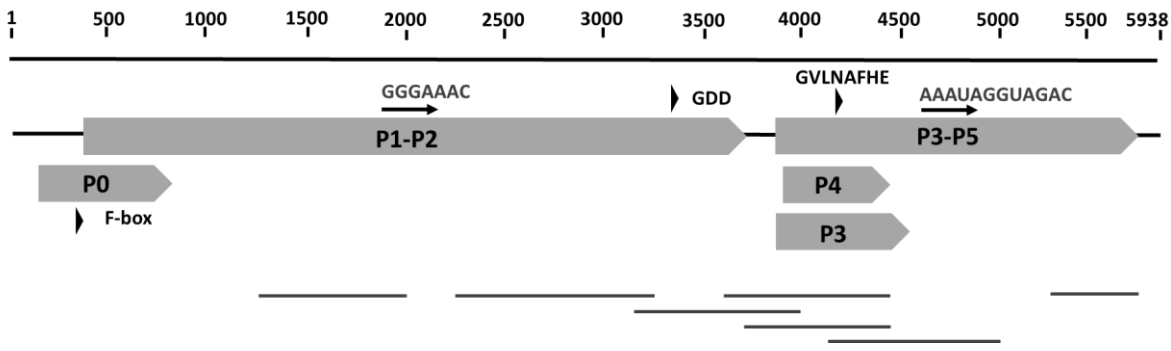


Fig. 1 Schematic representation and characterization of the genome organization of cassava polerovirus. ORFs are represented by a dark gray box indicating the viral sense (+) orientation. The putative pseudoknot starts with the GGGAAAC heptamer and the P5 protein translation after the stop codon UAG. Black triangles indicate poleroviruses conserved regions and gray lines below the sequence indicate regions confirmed by Sanger sequencing.

Phylogenetic analysis showed that the cassava polerovirus sequence (AJB1 S86) appears to be a distinct polerovirus that does not cluster with the other polerovirus sequences (Fig.2).

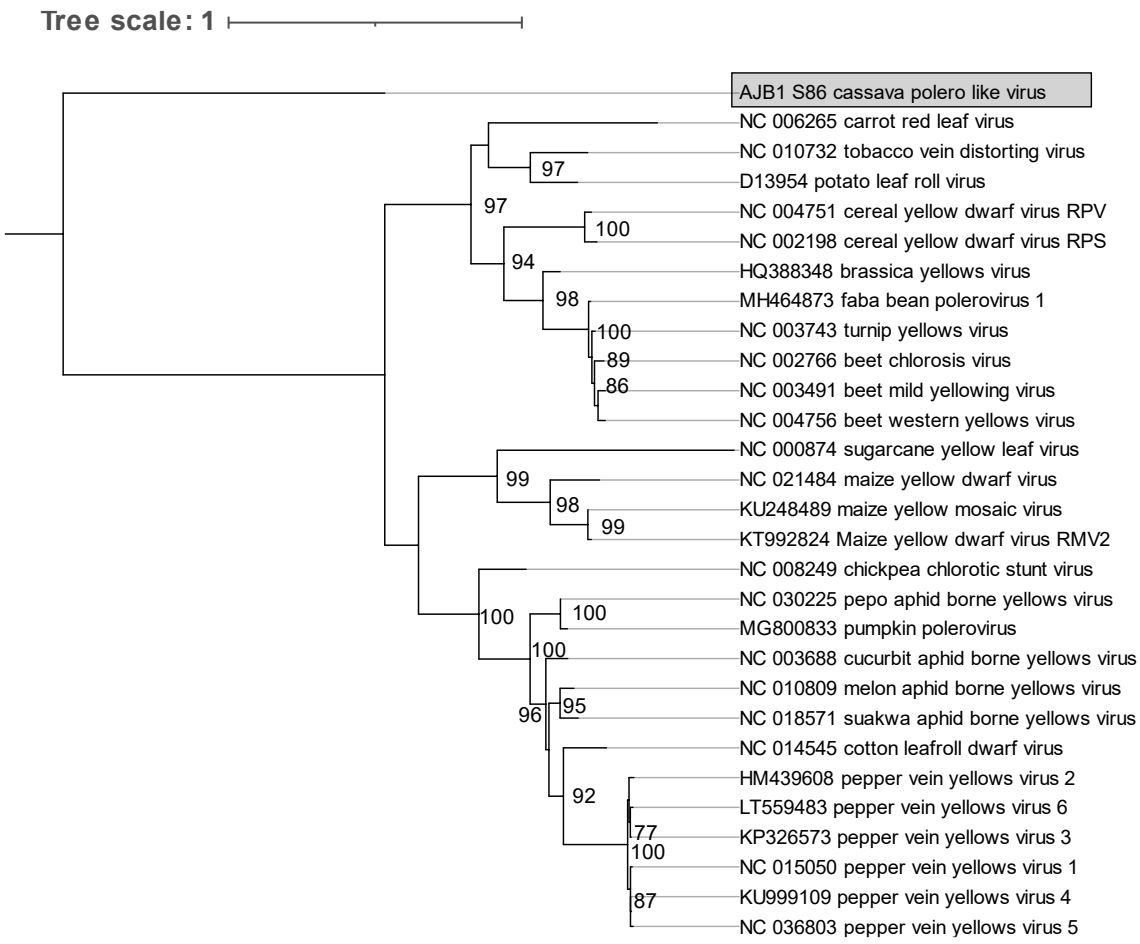


Fig. 2 Maximum likelihood tree reconstruction of poleroviruses based on fused proteins P3-P5. Bootstraps values are displayed above 75 at nodes. P3P5 aa sequences of poleroviruses available at National Center for Biotechnology Information (NCBI) were employed: beet chlorosis virus (NC_002766), beet mild yellowing virus (NC_003491), beet western yellows virus (NC_004756), carrot red leaf virus (NC_006265), cereal yellow dwarf virus-RPS (NC_002198), cereal yellow dwarf virus-RPV (NC_004751), chickpea chlorotic stunt virus (NC_008249), cotton leafroll dwarf virus (NC_014545), cucurbit aphid-borne yellows virus (NC_003688), faba bean polerovirus 1 (MH464873), maize yellow dwarf virus-RMV (NC_021484), maize yellow mosaic virus (KU248489), maize yellow mosaic virus (KT992824), melon aphid-borne yellows virus (NC_010809), pepo aphid-borne yellows virus (NC_030225), pepper vein yellows virus 1 (NC_015050), pepper vein yellows virus 2 (HM439608), pepper vein yellows virus 3 (KP326573), pepper vein yellows virus 4 (KU999109), pepper vein yellows virus 5 (NC_036803), pepper vein yellows virus 6 (LT559483), potato leafroll virus (D13954), pumpkin polerovirus (MG800833), suakwa aphid-borne yellows virus (NC_018571), sugarcane yellow leaf virus (NC_000874), tobacco vein distorting virus (NC_010732), turnip yellows virus (NC_003743) and brassica yellows virus (HQ388348).

Blastp analysis using the fused protein P1P2 sequence showed 95.06% identity (AHA91815) and 97.35% for the P3P5 fusion protein region (ALX37930), with partial sequences from Colombia. Distance matrix of SDT analysis (pairwise nt and aa alignment) using the phylogenetically closest polerovirus sequences showed that the cassava pole-like virus is a distinct species of polerovirus, since the highest identity of 47,2% (nt) and 47,9(aa) were found with the PLRV CP nt and TuYV P1P2 sequences, respectively (Supplementary Table 2). Our results indicate that the CsPLV is a distinct member of the *Polerovirus* genus.

Cassava polero-like virus was first reported infecting cassava in 2014 (Carvajal-Yepes et al. 2014). This virus was found with viruses of the *Secoviridae* and *Alphaflexiviridae* families in cassava plants expressing FDS (Oliveira et al. 2020; Carvajal-Yepes et al. 2014). However, there is no evidence that proves the viral association with the FSD, future studies are needed to understand the importance of this virus in cassava crops in different regions of Brazil.

This work shows the presence of a polerovirus infecting cassava plants in the Amazon region and shed light on the viral composition in cassava plantations. Additionally, in this work was assembled and characterized the first complete genome of CsPLV of the world.

Acknowledgements

We thank the Brazilian Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for the postgraduate Doctor's degree fellowship to the first author. This work was supported by Rural Federal University of Pernambuco (Edital 015/2018-UFRPE PRPPG).

Author information

Affiliations

Universidade Federal Rural de Pernambuco, Recife, Pernambuco, Brazil.
Géssyka Rodrigues de Albuquerque & Rosana Blawid
Embrapa Amazônia Oriental, Belém, Pará, Brazil.
Alessandra de Jesus Boari, Elisa Ferreira Moura Cunha & Késsia Fátima Cunha Pantoja.

Corresponding author

Correspondence to Rosana Blawid.

References

Abraham A, Vetten HJ (2022) Chickpea chlorotic stunt virus: a threat to cool-season food legumes. Archives of Virology 167: 21-30

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko I, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: A New Genome Assembly Algo-rithm and Its Applications to Single-Cell Sequencing. Journal of Computational Biology 19: 455-477

Blawid R, Silva JMF, Nagata T (2017) Discovering and sequencing new plant viral genomes by next-generation sequencing: description of a practical pipeline. Annals of Applied Biology 170: 301-314

Byrne MJ, Steele JFC, Hesketh EL, Walden M, Thompson RF, Lomonossoff GP, Ranson, NA (2019) Combining Transient Expression and Cryo-EM to Obtain HighResolution Structures of Luteovirid Particles. Structure 27, 1761-1770.e3.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30: 2114-2120

Carvajal-Yepes M, Olaya C, Lozano I, Cuervo M, Castaño M, Cuellar WJ (2014) Unraveling complex viral infections in cassava (*Manihot esculenta* Crantz) from Colombia. Virus Research 186:76–86

Cilia M, Peter KA, Bereman MS, Howe K, Fish T, Smith D, Gildow F, MacCoss MJ, Thannhauser TW, Gray SM (2012) Discovery and Targeted LC-MS/MS of Purified Polerovirus Reveals Differences in the Virus-Host Interactome Associated with Altered Aphid Transmission. PLoS ONE 7: e48177.

Costa TM, Blawid R, Aranda MA, Freitas D, Andrade GP, Inoue-Nagata AK, Nagata T (2019). Cucurbit aphid-borne yellows virus from melon plants in Brazil is an interspecific recombinant. Archives of virology 164: 249–254.

Darriba D, Taboada GL, Doallo R, Posada D (2011) ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics 8:1164-1165

Darty K, Denise A, Ponty Y (2009) VARNA: Interactive drawing and editing of the RNA secondary structure. Bioinformatics, 25:1974–1975.

Delfosse VC, Barrios Barón MP, Distéfano AJ (2021) What we know about poleroviruses: advances in understanding the functions of polerovirus proteins. Plant Pathololgy 70:1047-1061

Distéfano AJ, Bonacic Kresic I, Hopp HE (2010) The complete genome sequence of a virus associated with cotton blue disease, cotton leafroll dwarf virus, confirms that it is a new member of the genus Polerovirus. Archives of Virology 155:1849-1854.

Ghosh S, Kanakala S, Lebedev G, Kontsedalov S, Silverman D, Alon T, Mor N, Sela N, Luria N, Dombrovsky A, Mawassi M, Haviv S, Czosnek H, Ghanim M (2019) Transmission of a new polerovirus infecting pepper by the whitefly Bemisia tabaci. Journal of Virology 93:e00488–e00519.

Gray S, Gildow FE (2003) Luteovirus-aphid interactions. Annual Review of Phytopathology 41:539-566.

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. Systematic Biology, 59:307-321.

Jaag HM, Kawchuk L, Rohde W, Fischer R, Emans N, Prüfer D (2003). An unusual internal ribosomal entry site of inverted symmetry directs expression of a potato leafroll polerovirus

replication-associated protein. Proceedings of the National Academy of Sciences of the United States of America, 100:8939-8944.

Jeevalatha A, Kaundal P, Shandil RK, Sharma NN, Chakrabarti SK, Singh BP (2013) Complete Genome Sequence of Potato leafroll virus Isolates Infecting Potato in the Different Geographical Areas of India Shows Low Level Genetic Diversity. Indian journal of virology 24:199-204.

Letunic I, Bork P (2021) Interactive Tree of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Research 49:W293–W296

Letunic I, Khedkar S, Bork P (2021) SMART: recent updates, new developments and status in 2020. Nucleic Acids Research 49:D458–D460

Muhire BM, Varsani A, Martin DP (2014) SDT: A Virus Classification Tool Based on Pairwise Sequence Alignment and Identity Calculation. PLoS ONE 9:e108277

Oliveira SAS, Ferreira CF, Diamantino MSAS, Santos TA, Pereira JD, Oliveira EJ (2020) First report of cassava torrado-like virus, cassava polero-like virus and cassava new alphaflexivirus associated with cassava frogskin disease in Brazil. Journal of Plant Pathology. 10: 247.

Schwarz G (1978) Estimating the Dimension of a Model. Annals of Statistics 6:461-464

Smirnova E, Firth AE, Miller WA, Scheidecker D, Brault V, Reinbold C, Rakotondrafara AM, Chung BYW, Ziegler-Graff V (2015) Discovery of a Small Non-AUG-Initiated ORF in Poleroviruses and Luteoviruses That Is Required for Long-Distance Movement. PLoS Pathogens 11: e1004868

Sõmera M, Fargette D, Hébrard E, Sarmiento C (2021). ICTV Virus Taxonomy Profile: Solemoviridae 2021. The Journal of general virology, 102:001707.

Kumar S, Stecher G, Li M, Knyaz C, and Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Molecular Biology and Evolution 35:1547-1549

Zhao K, Yin Y, Hua M, Wang S, Mo X, Yuan E, Zheng H, Lin L, Chen J, Lu Y, Chen H, Lu Y, Chen J, Peng J, Yan F (2021) Pod pepper vein yellows virus, a new recombinant polerovirus infecting *Capsicum frutescens* in Yunnan province, China. Virology Journal 18:42

First report of a polerovirus infecting cassava plants from the Amazon region.

[1]Géssyka Rodrigues de Albuquerque, [2]Alessandra de Jesus Boari [2]Elisa Ferreira Moura Cunha, [2]Késsia Fátima Cunha Pantoja, [1]Rosana Blawid

1 Universidade Federal Rural de Pernambuco, Recife, Pernambuco, Brazil.
2 Embrapa Amazônia Oriental, Belém, Pará, Brazil.

Supplementary Table 1. List of primers used in this study for sanger sequence and detection of the cassava polerovirus.

Name	Length	Direction	Sequence (5'-3')	Amplicon (bp)
317	22	F	CGGAGAGACTGCATTGAACACA	3028
3343	20	R	CATAGCGACGCACCAAGAAG	
1148	20	F	AGCCGTGCATTATGTCACCT	1163
2309	20	R	TGAGTTTTTGGCCCCAGAGG	
2290	20	F	CCTCTGGGGCCAAAAACTCA	1054
3343	20	R	CATAGCGACGCACCAAGAAG	
3243	21	F	CTGGAGTCCAAAAGTCTGGCT	1001
4243	20	R	CCGCGAGTGACTGATACTGT	
3754	23	F	TGAAAGACGTATCCACTTACAGG	937
4708	23	R	AAGTCATTTCAGGCGGTACTTTA	
4204	20	F	AAGATCACAAGGCTCCAGGT	1000
5203	25	R	CTTTGTCTGTATAAGCGCTAGAACC	
5089	22	F	TGGTATACGTTGGCTATTGGGG	557
5644	22	R	AGAGTTCCTCCTGCCAAAGATG	

First report of a polerovirus infecting cassava plants from the Amazon region.

[1]Géssyka Rodrigues de Albuquerque, [2]Alessandra de Jesus Boari [2]Elisa Ferreira Moura Cunha, [2]Késsia Fátima Cunha Pantoja, [1]Rosana Blawid

1 Universidade Federal Rural de Pernambuco, Recife, Pernambuco, Brazil.

2 Embrapa Amazônia Oriental, Belém, Pará, Brazil.

Supplementary Table 2. List of primers used for race amplification.

Name	Length	Direction	Sequence (5'-3')	Description (3'-UTR)
4204	20	F	AAGATCACAAGGCTCCAGGT	cDNA
5089	22	F	TGGTATACGTTGGCTATTGGGG	First Nested-PCR
5415	22	F	TTGGAAGAATGTTGAACCAGGC	Second Nested-PCR
Name	**Length**	**Direction**	**Sequence (5'-3')**	**Description (5'-UTR)**
1167	20	R	AGGTGACATAATGCACGGCT	cDNA
890	23	R	CTGTGGGATAGGGTATGAGTTGA	First Nested-PCR
400	20	R	ATTGCCATCGGTTTGCATCG	Second Nested-PCR

First report of a polerovirus infecting cassava plants from the Amazon region.

[1]Géssyka Rodrigues de Albuquerque, [2]Alessandra de Jesus Boari [2]Elisa Ferreira Moura Cunha, [2]Késsia Fátima Cunha Pantoja, [1]Rosana Blawid

1 Universidade Federal Rural de Pernambuco, Recife, Pernambuco, Brazil.

2 Embrapa Amazônia Oriental, Belém, Pará, Brazil.

Supplementary Table 2. Distance matrix of SDT analysis (pairwise nt and aa alignment) of the closest polerovirus species phylogenetically, using the cassava polero-like virus (AJB1_S86) as reference.

Virus *		BChV	BMYV	BWYV	BrYV	CtRLV	CYDV-RPS	CYDV-RPV	FBPV1	PLRV	TVDV	TuYV
Accession no.		NC_002766	NC_003491	NC_004756	HQ388348	NC_006265	NC_002198	NC_004751	MH464873	D13954	NC_010732	NC_003743
nt	P0	33,3	30,8	33,0	32,7	31,7	32,1	32,2	32,9	32,1	32,0	32,6
	P1P2	41,2	41,5	41,7	43,2	40,4	40,5	40,0	41,8	41,4	41,6	43,1
	P3P5	40,7	41,3	42,3	42,0	40,1	41,1	41,1	43,9	42,7	41,3	41,5
	MP	42,4	42,6	42,0	43,2	44,1	45,4	43,9	43,4	46,4	45,7	43,1
	CP	42,7	43,9	42,8	42,8	42,7	45,5	43,3	43,1	47,2	45,9	42,7
aa	P0	25,7	26,8	30,3	26,8	24,7	21,5	24,4	25,2	25	22	25,4
	P1P2	42,3	44,4	44,9	47,7	45,0	45,5	44,8	46,1	42,3	45,0	47,9
	P3P5	33,6	31,0	32,0	31,5	31,4	31,4	33,2	33,8	33,1	31,1	33,1
	MP	30,3	28,1	30,8	28,2	33,3	37,3	33,8	30,8	35,6	36,4	29,7
	CP	41,2	40,7	41,9	40,5	41,8	46,9	46,0	42,4	45,6	40,7	42,2

* beet_chlorosis_virus (BChV), beet_mild_yellowing_virus (BMYV), beet_western_yellows_virus (BWYV), brassica_yellows_virus (BrYV), carrot_red_leaf_virus (CtRLV), cereal_yellow_dwarf_virus_RPS (CYDV-RPS), cereal_yellow_dwarf_virus_RPV (CYDV-RPV),

faba_bean_polerovirus_1 (FBPV1), potato_leaf_roll_virus (PLRV), tobacco_vein_distorting_virus (TVDV) and turnip_yellows_virus (TuYV).

CAPÍTULO V

Conclusões Gerais

CONCLUSÕES GERAIS

1. A análise de dados de 28 amostras de HTS de plantas de mandioca da Embrapa Mandioca e Fruticultura da Bahia, Cruz das Almas, sequenciadas no Departamento de Vírus de Plantas do Instituto DSMZ, na Alemanha, contribuiu para elucidação da diversidade viral no Brasil com a descoberta de uma nova espécie do gênero *Torradovirus* que foi nomeada de cassava Torrado-like virus 2 (CsTLV-2). Foi possível montar o primeiro genoma completo da nova espécie e desenhar primers de detecção para o uso em RT-PCR.

2. O CsTLV-2 também foi detectado em uma amostra de mandioca da coleção da Embrapa Amazônia Oriental e foi possível observar que havia diversidade genética entre os isolados do Norte e Nordeste do Brasil. Primers de detecção foram desenhados e puderam ser utilizados para amplificar sequências dos dois isolados.

3. A análise do HTS de um pool de 14 amostras de mandioca da coleção da Embrapa Amazônia Oriental, mostrou que além do CsTLV-2, havia mais duas espécies de vírus na região. Uma delas, o CsVX, que ainda não havia sido detectada no Brasil e o CsPLV. O primeiro genoma quase completo de CsVX do Brasil foi montado e foi possível desenhar e validar primers de detecção para o isolado brasileiro. Na mesma amostra, também foi possível detectar um satélite associado aos potexvírus. Assim como, foi possível montar e validar parcialmente por Sanger o primeiro genoma quase completo de CsPLV.

4. O uso do HTS foi essencial para a descoberta de novos vírus no Brasil e o tamanho da amostragem de plantas de mandioca das regiões Norte e Nordeste realizada neste trabalho foi considerada importante para fornecer um panorama da diversidade viral nessas regiões. A descoberta e o estabelecimento da detecção desses vírus geraram informações importantes que contribuirão com o estabelecimento de medidas de controle para prevenir a dispersão e trânsito de material vegetal propagativo infectado e com os programas de melhoramento da Mandioca no Brasil.

ANEXO I

Tabela 1. Resultado do tBlastx dos acessos de HTSs analisados e dados de montagem.

Código do HTS	Código GenBank	% Id	Descrição	E Value	Max Seq*	Cobertura	n Reads
DSC407_S20	MF449522	96.6%	Cassava Torrado-like virus isolate Sec13 polyprotein	0	2377	145	2539
	MF449523	97.0%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	0	1202	254	3053
DSC409_S2	MF449522	96.0%	Cassava Torrado-like virus isolate Sec13 polyprotein	0	2059	16	197
	MF449523	96.4%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	0	1231	61	370
	NC_033817	91.3%	Cassava satellite virus isolate Casatv_Br	0	264	12	29
DSC416_S19	MF449522	97.2%	Cassava Torrado-like virus isolate Sec13 polyprotein	0	1615	39	371
	MF449523	95.8%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	0	1247		
DSC417_S8	NC_001648	95.4%	Cassava vein mosaic virus	0	892	32	209
DSC430_S11	NC_016136	60.4%	Potato yellow dwarf virus	0	879	36	617
	NC_016136	42.6%	Potato yellow dwarf virus	0	298	37	415
P25_6A_S9	NC_033739	67.3%	Cacao yellow vein-banding virus isolate ICS27	1.25e-119	251		
	KR108830	57.8%	Cassava brown streak virus isolate TZ_Ser_6 polyprotein	1.55e-58	161		
	HG965221	58.0%	Cassava brown streak virus isolate CBSV	7.83e-58	157		
DSCG7_S18	MF449522	95.9%	Cassava Torrado-like virus isolate Sec13 polyprotein	0	2026		
	MF449523	97.0%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	0	977		

DSC492_S12	MF449522	96.1%	Cassava Torrado-like virus isolate Sec13 polyprotein	0	2297	30	914
	MF449523	96.1%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	0	1249	105	1157
DSC442_S10	NC_025389	65.1%	Eggplant mottled dwarf virus isolate Agapanthus	2.08e-118	192		
	NC_016136	50.0%	Potato yellow dwarf virus	7.97e-48	148		
DSC437_S9	MF449523	96.6%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	0	1074		
	MF449522	96.8%	Cassava Torrado-like virus isolate Sec13 polyprotein	0	745		
	NC_009992	45.7%	Plum bark necrosis and stem pitting-associated virus	0	435		
DSC445_S7	NC_040841	44.5%	Pinus nigra virus 1 isolate B2	1.10e-60	191		
	NC_028462	23.3%	Blueberry fruit drop associated virus	3.37e-16	180		
	MN243731	90.3%	Cassava common mosaic virus isolate Hainan-DZ	1.82e-83	144		
DSC382_S6	NC_028462	26.9%	Blueberry fruit drop associated virus	2.26e-11	134		
	NC_040841	53.1%	Pinus nigra virus 1 isolate B2	2.14e-46	128		
DSC372_S5	MN243731	95.5%	Cassava common mosaic virus isolate Hainan-DZ	0	972	1076	34720
	NC_016436	36.2%	Grapevine leafroll-associated virus 7	1.13e-18	116		
P26_2E_S13	MF449522	96.1%	Cassava Torrado-like virus isolate Sec13 polyprotein	0	2382	1922	63587
	MF449523	96.0%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	0	1240	2821	115578
P26_4C_S11	NC_009992	40.6%	Plum bark necrosis and stem pitting-associated virus	0	503		
	MF449523	72.0%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	0	418	60	606

	MF449522	66.9%	Cassava Torrado-like virus isolate Sec13 polyprotein	0	354	22	292
DSC416_S8	MF449523	96.3%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	0	409		
	MF449522	94.5%	Cassava Torrado-like virus isolate Sec13 polyprotein	0	289		
DSC369_S4	MF449522	95.7%	Cassava Torrado-like virus isolate Sec13 polyprotein	0	1952		
	MF449523	95.6%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	0	1251		
DSC407_S9	MF449523	95.7%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	0	1244	36	352
	MF449522	97.2%	Cassava Torrado-like virus isolate Sec13 polyprotein	0	604		
DSC435_S10	NC_040841	45.5%	Pinus nigra virus 1 isolate B2	1.03e-69	143		
DSC389_S3	NC_001648	94.0%	Cassava vein mosaic virus	0	798	643	7516
DSCG6_S6	MF449523	87.2%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	0	337		
	NC_001497	47.5%	Cauliflower mosaic virus	2.70e-51	179		
	MF449522	83.1%	Cassava Torrado-like virus isolate Sec13 polyprotein	8.79e-82	142		
DSCG3_S3	NC_033817	94.1%	Cassava satellite virus isolate Casatv_Br	7.74e-144	239		
	NC_009992	52.0%	Plum bark necrosis and stem pitting-associated virus	1.71e-54	173		
DSCG2_S2	MF449522	96.6%	Cassava Torrado-like virus isolate Sec13 polyprotein	0	1151		
	MF449523	97.3%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	0	1145	50	466
	NC_001648	92.6%	Cassava vein mosaic virus	0	470		

	NC_003674	70.7%	Olive latent virus 2 RNA 2	4.07e-167	317		
	NC_003673	70.2%	Olive latent virus 2 RNA 1	0	248		
P27_6F_S1	NC_009992	40.6%	Plum bark necrosis and stem pitting-associated virus	0	503	104	2482
	MF449523	72.4%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	0	413	1630	25367
	MF449522	66.9%	Cassava Torrado-like virus isolate Sec13 polyprotein	0	354	790	23436
P27_6G_S2	NC_009992	40.6%	Plum bark necrosis and stem pitting-associated virus	0	503	154	2786
	MF449523	72.4%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	0	413	2814	56671
	MF449522	69.5%	Cassava Torrado-like virus isolate Sec13 polyprotein	0	315	1309	40354
P27_4C_W44	NC_003549	98.2%	Cowpea mosaic virus RNA 1	0	1558		
	NC_003550	98.1%	Cowpea mosaic virus (CPMV) middle-component RNA (M RNA)	0	933		
	NC_009992	40.6%	Plum bark necrosis and stem pitting-associated virus	0	503	244	9677
	MF449523	72.0%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	0	418	1179	19682
	MF449522	66.9%	Cassava Torrado-like virus isolate Sec13 polyprotein	0	354	372	1754
DSCG1_S1							
DSCG12_S12	MN243731	75.8%	Cassava common mosaic virus isolate Hainan-DZ	0	1341		
	MF449522	98.9%	Cassava Torrado-like virus isolate Sec13 polyprotein	0	353		

	NC_025389	39.0%	Eggplant mottled dwarf virus isolate Agapanthus, complete genome	5.40e-74	272		
	MF449523	100.0%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	8.37e-165	237		
DSC404_S1							
DSCG5_S5	MF449522	97.2%	Cassava Torrado-like virus isolate Sec13 polyprotein	0	1406		
	MF449523	96.6%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	0	1212		
DSCG7_S7	MF449523	85.0%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	0	601		
	MF449522	94.3%	Cassava Torrado-like virus isolate Sec13 polyprotein	3.24e-135	210		
DSCG4_S4	MF449522	97.5%	Cassava Torrado-like virus isolate Sec13 polyprotein	0	1867	68	840
	MF449523	95.9%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	0	1248	107	1090
	NC_001648	91.4%	Cassava vein mosaic virus	0	627		
	NC_003674	69.9%	Olive latent virus 2 RNA 2	0	508		
	NC_003673	57.4%	Olive latent virus 2 RNA 1	0	394		
SRR1050897	MF449522	95.6%	Cassava Torrado-like virus isolate Sec13 polyprotein	0	2400	256	6059
	MF449523	95.8%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	0	1244	598	28337
	NC_034375	93.1%	Cassava virus X isolate Ven164	0	922	349	6953
SRR10480878	NC_034375	95.6%	Cassava virus X isolate Ven164	0	893	50	756
	MN243731	98.0%	Cassava common mosaic virus	0	745	43	665
AJB1_S86	NC_034375	83.5%	Cassava virus X isolate Ven164	0	1083	148	1747
	KC505249	87.0%	Cassava Polero-like virus	0	5938	511	9971

MF449522	53.2%	Cassava Torrado-like virus isolate Sec13 polyprotein	0	310	50	970
MF449523	69.5%	Cassava Torrado-like virus isolate Sec13 hypothetical protein and polyprotein	7.95e-152	308	25	2520
NC_033817	89.6%	Cassava satellite virus	0	307	361	1623