Universidade de Brasília
Instituto de Ciências Biológicas
Departamento de Fitopatologia
Programa de Pós-Graduação em Fitopatologia

High Throughput Sequencing (HTS)–based identification and characterization of *Geminiviridae* family members in weeds associated with the tomato crop in Brazil.

EDUARDO SOARES DA SILVA LIMA

Brasília–DF
2023

Eduardo Soares da Silva Lima

High Throughput Sequencing (HTS)–based identification and characterization of *Geminiviridae* family members in weeds associated with the tomato crop in Brazil.

Dissertation presented to the University of Brasília as a partial requirement for obtaining the title of Master in Phytopathology by the Post-Graduate Program in Phytopathology.

Advisor

Profª Drª. Rita de Cássia Pereira Carvalho

Brasília–DF
2023

CATALOG SHEET

High Throughput Sequencing (HTS)–based identification and characterization of *Geminiviridae* family members in weeds associated with the tomato crop in Brazil.

Brasília-DF, 2023

Pages number p.: 97.

Master Thesis – Graduate Program in Phytopathology, University of Brasilia, Brasília, DF.

I- *Begomovirus*, ssDNA, satellite DNAs

II- University of Brasilia. Graduate Program in Phytopathology – PPG/FIT.

A Deus primeiramente, pela vida, por estar sempre dando graças, cooperando, iluminando os meus caminhos. A minha família, principalmente minha mãe pelo exemplo de luta, de força e de amor

Dedicatória

Agradecimentos

A capacidade de sonhar sempre foi o grande segredo daqueles que mudaram o mundo. Os sonhos alimentam a alma e dão asas a inteligência. É no solo fértil da memória onde semeamos os sonhos que farão grande diferença em nossa existência. Os sonhadores mudaram a história da humanidade. Eles fizeram da derrota, o pódio para a vitória; das críticas, o palco, de onde receberam os aplausos (Lihat Sedikit). Sonho em viver com saúde, em ter comigo a minha família até envelhecer, ter estabilidade financeira, fatores que me movem a continuar. Por meio deste quero vir agradecer a tudo e todos àqueles que estão a me ajudar a concretizar e estimular novos sonhos.

Quero agradecer a Universidade de Brasília (UnB) por proporcionar a realização deste curso;

Ao Programa de Pós-Graduação em Fitopatologia da UnB;

Aos professores que se dedicaram e compartilharam seus conhecimentos;

A minha orientadora, Profª. Drª. Rita de Cássia Pereira-Carvalho, quero agradecê-la pela paciência, compreensão, entusiasmo e confiança. Por ter sido um exemplo de pessoa, educadora, pesquisadora e orientadora.

Ao Dr. Leonardo Boiteux e Dra. Maria Esther Boiteux pela dedicação e sempre ajudarem compartilhando seus conhecimentos.

Ao técnico do Laboratório de Genômica e Melhoramento da Embrapa Hortaliças, Antônio Francisco Costa, sempre disposto a ajudar.

Aos membros do Laboratório de Virologia Vegetal, amigos, Izaías sempre disposto a ajudar, Henrique por compartilhar seus conhecimentos e conversas, Marcos por compartilhar seus conhecimentos e entusiasmo, Giovana, Priscila, Maria Luiza, Amanda e Juliana, por todo apoio, torcida e alegrias compartilhadas. Ao Felipe com apoio técnico, pela paciência, risadas e agora amigo. A Drª Luciane Reis, por compartilhar seus conhecimentos, pela paciência e sempre disposta a ajudar.

Aos amigos que a Fitopatologia me trouxe e que levarei para a vida, Erivaldo, Ian, Luciellen, Angélica, Alice, Isabela, Jefferson, Tais e Tiago.

A minha família, minha mãe, Francisca como exemplo de força e persistência, as minhas irmãs, Karla e Andréa por me apoiarem, me fazerem rir e sempre estarem presentes quando preciso. Pelas minhas sobrinhas, Paula, Anne, Anna Julia e Patrícia, que me trazem alegrias e apoio.

Ao meu esposo pela paciência e suporte para me ajudar, para que eu conseguisse trilhar da melhor forma.

Work carried out in the Department of Plant Pathology of the Institute of Biological Sciences of the University of Brasília (UnB), under the guidance of Dr. Rita de Cássia Pereira Carvalho.

High Throughput Sequencing (HTS)–based identification and characterization of *Geminiviridae* family members in weeds associated with the tomato crop in Brazil.

EDUARDO SOARES DA SILVA LIMA

Thesis approved 18/04/2023 by:

Dra. Rita de Cássia Pereira-Carvalho
Universidade de Brasília – UnB
(Presidente da Banca)

Prof°. Dr. Robert G. Miller
Departamento de Biologia Celular – Universidade de Brasília – UnB
(Membro Interno da Banca)

Dra. Mirtes Freitas Lima
Embrapa Hortaliças
(Membro Externo da Banca)

Prof. Helson Mario Martins do Vale
Departamento de Fitopatologia – Universidade de Brasília – UnB
(Membro Suplente)

SUMMARY

LIST OF FIGURES 9
LIST OF TABLES 10
RESUMO GERAL 11
GENERAL ABSTRACT 13
GENERAL INTRODUCTION 15
1.1. Tomato crop..19
1.2 Viral pathogens in tomato crops..20
1.3 *Geminiviridae* 211
1.4 *Begomovirus* 23
1.5 DNA satellite 25
1.6 *Begomovirus*-host range 25
1.7 *High-Throughput Sequencing* (HTS) 26
REFERENCES 27
CHAPTER 2. Circular, single-stranded DNA virome of neotropical weeds associated with tomato fields..37
CHAPTER 3. Complete genome characterization of a novel recombinant *Begomovirus* species infecting a new natural Fabaceae host (*Bolusafra bituminosa)* 70

LIST OF FIGURES

CHAPTER 1. LITERATURE REVIEW

Figure 1- Estimates of tomato production, according to the main producing Federation Units and annual variation (%) - Brazil – 1994 and 2021...14

Figure 2. Genomic organization of *Begomovirus*. ……………………………………17

Figure 3. Introducing vegetables with typical symptoms of begomoviroses…………26

CHAPTER 2. CIRCULAR, SINGLE-STRANDED DNA VIROME OF NEOTROPICAL WEEDS ASSOCIATED WITH TOMATO FIELDS

Figure 1. Circular, single-stranded DNA virome of neotropical weeds associated with tomato fields...54

Figure 2. Middle root Bayesian phylogenetic tree with GTR+I model) for DNA–A components of the selected begomoviruses. 56

CHAPTER 3. COMPLETE GENOME CHARACTERIZATION OF A NOVEL BIPARTITE BEGOMOVIRUS SPECIES INFECTING A NEW NATURAL FABACEAE HOST (*Bolusafra bituminosa)*

Figure 1. Genomic organization of the CE-76 species named as *Bolusafra interveinal chlorosis virus*...77

Figure 2. Sequence Demarcation Tool (SDT) pair-by-pair identity analysis performed using DNA–A sequence information from *Begomovirus* species obtained from the NCBI database. Middle root Bayesian phylogenetic tree with GTR+I model) for DNA–A components Bolusafra interveinal chlorosis virus (BoICV)……………………………78

LIST OF TABLES

CHAPTER 1. LITERATURE REVIEW

Table 1. List of genera of the *Geminiviridae* family considering the number of species by genus, host group, vector, genomic organization and references....................................22

CHAPTER 2. CIRCULAR, SINGLE-STRANDED DNA VIROME OF NEOTROPICAL WEEDS ASSOCIATED WITH TOMATO FIELDS

Table 1. Number of contigs, reads coverage, assembled genome size, BLAST coverage, sequence identity of the assembled virus, E-value, access to viruses and subviral agents obtained by High-Throughput Sequencing (HTS) for the DNA-A segment and description of the pool containing 94 begomovirus samples.....................................49

Table 2. Number of contigs, read coverage, assembled genome size, BLAST coverage, assembled virus sequence identity, E-value, access to viruses and subviral agents obtained by High-Throughput Sequencing (HTS) for the DNA-B segment and description of the pool containing 94 begomovirus samples...51

Table 3. Number of contigs, read coverage, assembled genome size, BLAST coverage, assembled virus sequence identity, E-value, access to viruses and subviral agents obtained by High-Throughput Sequencing (HTS) for the DNA segment-A and subviral agents and description of the pool containing 06 samples of *Clecrusatellite*, *Mulcrilevirus* and, *Topilevirus*...53

Supplementary Table 1. Information about weeds classified in eleven botanical families associated with the crop tomato from in Brazil...64

CHAPTER 3. COMPLETE GENOME CHARACTERIZATION OF A NOVEL BIPARTITE BEGOMOVIRUS SPECIES INFECTING A NEW NATURAL FABACEAE HOST (*Bolusafra bituminosa*)

Supplementary Table 1. Begomoviruses reported in hosts belonging to the *Fabaceae* family...80

GENERAL ABSTRACT

Lima, Eduardo Soares da Silva. University of Brasilia, Brasilia-DF. High High Throughput Sequencing (HTS)–based identification and characterization of *Geminiviridae* family members in weeds associated with the tomato crop in Brazil. 2023. 97 pages. Dissertation (Master in Phytopathology). Advisor: Dr. Rita de Cássia Pereira-Carvalho.

The tomato (*Solanum lycopersicum* L.) crop has great economic and social importance for Brazil. Production of this vegetable is affected by many diseases, including those induced by *Begomovirus* species (family *Geminiviridae*). The **Chapter 1** presents a review of these pathogens and the diseases they induce. In addition to tomato, begomoviruses have a wide range of hosts across different botanical families. Many weeds are natural hosts, playing an important role as reservoirs for this group of pathogens. Begomoviruses are transmitted by the *Bemisia tabaci* species complex, which are characterized by having a wide geographic distribution and being extremely polyphagous. These characteristics of the vectors increase the potential for mixed viral infections and for events of recombination and pseudo-recombination. Such events provide an increase in the variability of the viral populations, including the emergence of novel species and viral variants capable of overcoming resistance factors present in commercial varieties. In this context, monitoring viral diversity in tomato fields and in associated weeds is extremely important for effective management of these pathogens. Different molecular strategies have been employed for large-scale identification and characterization of begomoviruses, including High Throughput Sequencing (HTS) platforms. Herein, a broad metagenomics analysis of single-stranded, circular DNA viral species of the *Geminiviridae* family (especially from the genus *Begomovirus*) was conducted in weeds frequently occurring in tomato fields. Plants from the families Malvaceae, Euphorbiaceae, Amaranthaceae, Solanaceae, Asteraceae, Rubiaceae, Fabaceae, Lamiaceae, Cucurbitaceae, Convolvulaceae, and Brassicaceae were collected in field surveys carried out in all five Brazilian regions. In **Chapter 2**, 91 foliar samples of weeds exhibiting begomovirus-like symptoms (mosaic, mottled, chlorotic sectors, and dwarfism) were selected according to the year and collection site. Samples were collected in production areas and areas in the vicinity of tomato fields between 2003 and 2022. The samples were subjected to total DNA extraction, which was used as a template for performing Rolling Circle Amplification (RCA) assays. The selected samples were submitted to PCR tests with specific primers targeting conserved genomic regions of begomoviruses. The RCA pool was assembled with positive PCR samples and sent for sequencing on an Illumina Nova Seq 6000 platform aiming to obtain viral genomes. For botanical identification of the host species, the DNAs of a subgroup of weed plants were used as template in PCR assays using primers targeting the Rubisco and/or Maturase K barcoding genes. After the assembly and recovery of 100 contigs, 20 corresponded to 15 new species that are currently being characterized biologically and molecularly. In **Chapter 3,** the complete genome of a genetically distinct isolate representing a potential new species (denominated CE–076) was retrieved via HTS and validated via Sanger sequencing. This potential new virus displayed identity of 85.87% with tomato bright yellow mottle virus (ToBYMV) and it was detected infecting a novel Fabaceae host species – *Bolusafra bituminosa*. The information generated in the present dissertation can contribute to the establishment of management systems and generate information of interest on viral diversity for tomato breeding programs. Furthermore, the present work confirms the epidemiological importance the weed plants as reservoirs of viral species described infecting tomato as well as a potential role in the genetic evolution of populations of this group of pathogens in Brazil.

Keywords: Begomovirus, diversity, *Solanum lycopersicum* L., Sanger dideoxy sequencing, genetic resistance.

RESUMO GERAL

Lima, Eduardo Soares da Silva. Universidade de Brasília, Brasília, DF. Identificação e caracterização baseada em High Throughput Sequencing (HTS) de membros da família *Geminiviridae* em plantas daninhas associadas à cultura do tomateiro no Brasil. 2023. 97 páginas. Dissertação (Mestrado em Fitopatologia). Orientadora: Profª Drª. Rita de Cássia Pereira-Carvalho.

A cultura do tomateiro (*Solanum lycopersicum* L.) tem grande importância econômica e social para o Brasil. A produção dessa hortaliça é afetada por diversas doenças, inclusive as induzidas por espécies de *Begomovirus* (família *Geminiviridae*). O **Capítulo 1** apresenta uma revisão desses patógenos e das doenças que eles induzem. Além do tomateiro, os begomovírus têm uma ampla variedade de hospedeiros em diferentes famílias botânicas. Muitas plantas daninhas são hospedeiras naturais, desempenhando importante papel como reservatórios desse grupo de patógenos. Os begomovírus são transmitidos pelo complexo de espécies *Bemisia tabaci*, que se caracterizam por terem ampla distribuição geográfica e serem extremamente polífagos. Essas características dos vetores aumentam o potencial para infecções virais mistas e a ocorrência de eventos de recombinação e pseudo-recombinação. Tais eventos contribuem para o aumento da variabilidade genética das populações virais, incluindo o surgimento de novas espécies e de variantes virais capazes de superar os fatores de resistência presentes nas variedades comerciais. Nesse contexto, o monitoramento da diversidade viral nas lavouras de tomate e nas plantas daninhas associadas é de extrema importância para o manejo eficaz desses patógenos. Diferentes estratégias moleculares têm sido empregadas para identificação e caracterização em larga escala de begomovírus, incluindo plataformas de High Throughput Sequencing (HTS). Aqui, uma ampla análise metagenômica de espécies virais de DNA circular de fita simples da família *Geminiviridae* (especialmente do gênero *Begomovirus*) foi realizada em plantas daninhas que ocorrem frequentemente em plantações de tomate. Plantas das famílias Malvaceae, Euphorbiaceae, Amaranthaceae, Solanaceae, Asteraceae, Rubiaceae, Fabaceae, Lamiaceae, Cucurbitaceae, Convolvulaceae e Brassicaceae foram coletadas em levantamentos de campo realizados nas cinco regiões brasileiras. No **Capítulo 2**, foram selecionadas 91 amostras foliares de plantas daninhas com sintomas de begomovírus (mosaico, mosqueado, setores cloróticos e nanismo) de acordo com o ano e o local de coleta. As amostras foram coletadas em áreas de produção e áreas próximas a plantações de tomate entre 2003 e 2022. As amostras foram submetidas à extração de DNA total, que serviu de molde para a realização dos ensaios de Rolling Circle Amplification (RCA). As amostras selecionadas foram submetidas a testes de PCR com primers específicos visando amplificar regiões genômicas conservadas de begomovírus. O pool de RCA foi montado com amostras de PCR positivas e enviado para sequenciamento em uma plataforma Illumina Nova Seq 6000 com o objetivo de obter genomas virais. Para a identificação botânica das espécies hospedeiras, os DNAs de um subgrupo de plantas daninhas foram usados como molde em ensaios de PCR usando primers direcionados aos genes *barcoding* Rubisco e/ou Maturase K. Após a montagem e recuperação de 100 contigs, 20 corresponderam a 15 novas espécies que estão sendo caracterizadas biológica e molecularmente. No **Capítulo 3**, o genoma completo de um isolado geneticamente distinto representando uma nova espécie em potencial (denominado CE–076) foi recuperado via HTS e validado via sequenciamento de Sanger. Este potencial novo vírus exibiu uma identidade de 85,87% com o ToBYMV (tomato bright yellow mottle virus) e foi detectado infectando uma nova espécie hospedeira de Fabaceae – Bolusafra bituminosa. As informações geradas na presente dissertação podem contribuir para o estabelecimento de sistemas mais eficientes de manejo e gerar informações de interesse sobre diversidade viral para os programas de melhoramento do tomateiro. Além disso, o presente trabalho confirma a importância epidemiológica das plantas daninhas como reservatórios de espécies virais descritas infectando o tomateiro, bem como um potencial papel na evolução genética de populações desse grupo de patógenos no Brasil.

Palavras chaves: *Begomovirus*, diversidade, *Solanum lycopersicum* L., sequenciamento Sanger, resistência genética.

GENERAL INTRODUCTION

The tomato (*Solanum lycopersicum* L.) production in Brazil is currently around 3.6 million tons, placing the country in 10th position in terms of world ranking (FAOSTAT 2023). However, a wide array of diseases might affect tomato yield and quality worldwide (Lopes and Ávila 2005; Jones 2014; Ong et al., 2020). Among the diseases caused by viruses, species classified in the genus *Begomovirus* (family *Geminiviridae*) cause yield losses between 30 and 100% (Giordano et al., 2005; Lopes and Reis 2017). Begomovirus genomes have either one (DNA–A only) or two DNA (DNA–A and DNA–B) molecules, being referred to as monopartite begomoviruses and bipartite begomoviruses, respectively. In bipartite species the DNA molecules are encapsidated separately. Transmission is done by the aleyrodid insects of the *Bemisia tabaci* complex in a relationship known as non-propagative, with the exception for tomato yellow leaf curl virus (TYLCV) (He et al., 2020).

In Brazil, the introduction of the vector *B. tabaci* Middle East Asia Minor 1 (=biotype B) in the 1990s was responsible for the intensification of the outbreaks induced by begomoviruses. These pathogens subsequently spread to several regions of the country (Cotrim et al., 2007; Reis et al., 2020; Reis et al., 2021). In addition, the entry of the vector possibly contributed to the spread of tomato severe rugose virus (ToSRV) to common bean (*Phaseolus vulgaris*) (Macedo et al., 2017a), soybean (*Glycine max*) (Macedo et al. 2017b), *Nicandra physalodes* (Barreto et al., 2013), *Capsicum* spp. (Bezerra-Agasie et al., 2006), eggplant (*Solanum melongena*) (Moura et al., 2018), *Physalis angulata* (Fernandes-Acioli et al., 2011; Duarte et al., 2020) and also to weeds such as *Macroptilium erythroloma* (Batista et al. 2022) and *Oxalis latifolia* (Pereira-Silva et al., 2022). The main symptoms of begomovirus in tomato comprise mosaic, leaf chlorotic spots, leaf deformation and curling and reduction in fruit size.

High-throughput sequencing (HTS) is an important and very robust tool in detecting and monitoring the diversity of viruses in crop plants as well as prospecting weed species as potential sources of inoculum. HTS enables the detection of existing viruses as well as new virus species. DNA viruses haven been detected via HTS in monocotyledonous and dicotyledonous species (Richet et al., 2019), including tomatoes (Reis et al., 2020).

The monitoring of begomovirus is essential to ensure the yield and quality of susceptible agricultural crops. Early detection and diversity assessment might contribute to establishment of integrated techniques for more appropriate management to prevent spread of the virus (Hasiów-Jaroszewska et al., 2021). In addition, the surveillance of weeds associated with the tomato crops is necessary verifying that they can act as alternative hosts.

In this context, the objective of the present work was to identify and characterize the diversity of viruses, mainly begomoviruses, in weeds belonging to the families, Amaranthaceae, Asteraceae, Brassicaceae, Convolvulaceae, Cucurbitaceae, Euphorbiaceae, Fabaceae, Lamiaceae, Malvaceae, Rubiaceae and Solanaceae associated with tomato cultivation with using HTS. This information was generated from different sampled geographic regions as well as different invasive plants associated with the tomato plant.

Hypothesis

Begomoviruses generally show high genetic variability due to a combination of distinct factors such as the high efficiency and multiple hosts of the vector, their small genomes (prone to recombination and pseudo-recombination events). These features are contributing to the diversity and emergence of new begomoviruses in weeds as their transfer to the associated crops.

Objetive

The main objective of the present study was to characterize (via HTS) the diversity of the begomoviruses present in tomato-associated weeds classified in the families Malvaceae, Euphorbiaceae, Amaranthaceae, Solanaceae, Asteraceae, Rubiaceae, Fabaceae, Lamiaceae, Cucurbitaceae, Convolvulaceae and Brassicaceae.

Specific objectives

- Screen and characterize viruses of the *Geminiviridae* family and satellite DNAs occurring in weeds (from 11 botanical families) collected with begomovirus-like symptoms either within or in the vicinity of tomato fields.

- Carry out a complete characterization of potential new bipartite *Begomovirus* species in foliar samples of a Fabaceae weed associated with the tomato crop in the Northeast Brazil.

CHAPTER 1

Literature Review

1.1. The tomato crop

The tomato (*Solanum lycopersicum* L.) belongs to the order Solanales, family Solanaceae, genus *Solanum* (section *Lycopersicon*) (NCBI 2023). The genus *Solanum* comprises about 1500 species, including species of economic importance, such potatoes, eggplant and tomatoes. *S. lycopersicum* originates from the Andes in South American and Galapagos Island, and north of Chile, Peru and Ecuador (Pereira-Carvalho et al., 2014). The domestication of the tomato occurred in Puebla and Vera Cruz (Mexico) (Camargo et al., 2006).

Tomato is considered one of the most important vegetable crops globally. China is currently the largest producer in the world and Brazil ranks ninth in production (FAOSTAT 2023) (**Figure 1**). Tomato is a vegetable crop adapted to tropical and subtropical regions. Mild temperatures (in the range of 18 to 25°C) favor germination and vegetative growth. Above these temperatures the flowers might drop, and the fruits become small and not suitable for commercialization.

The tomato is cultivated in all major Brazilian regions and practically all year round. This intensive cultivation provides favorable conditions for a wide array of pathogens, including begomovirus. According to the Systematic Survey of Agricultural Production in Brazil, for 2022, Goiás State was the largest producer, corresponding to 27.45 % of the total. In second position, the state of São Paulo (SP) produced 24.21 % and in third position the state of Minas Gerais (MG) with 14.22 %. The tomato harvested area in Brazil corresponds to 54,267 hectares (ha) with a total production of 3,886.009 tons in the year 2021. The average yield was 71,609 Kg/ha (IBGE 2022).

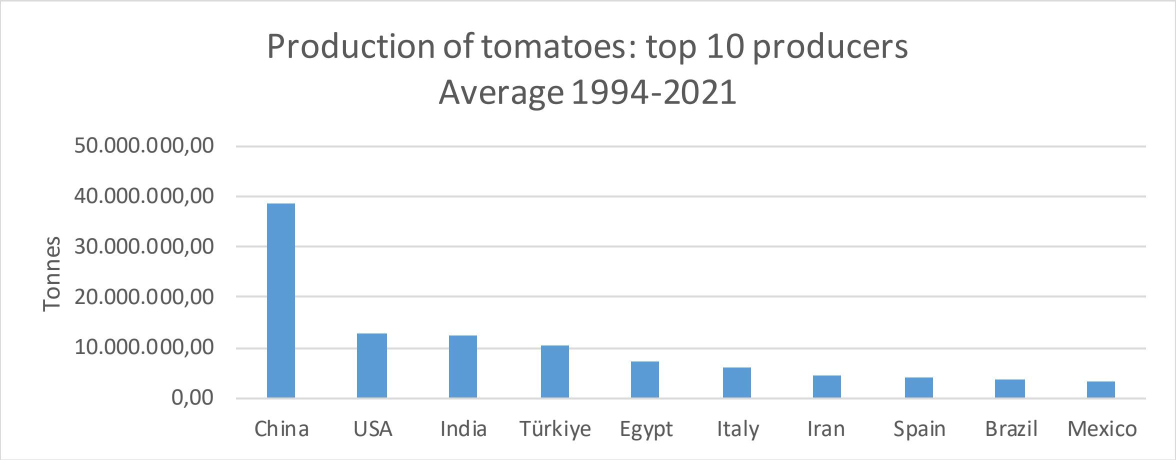


Figure 1. Top ten tomato-producing countries in the world (FAOSTAT 2023)

1.2 Tomato pathogens and diseases

Tomato production is of great economic and social importance. However, many diseases caused by viruses, fungi, oomycetes, bacteria, and nematodes may affect this crop (for review see Lopes and Ávila 2005; Lopes and Reis, 2011; Jones et al., 2014; Pinheiro et al., 2014).

1.3 Viral pathogens of tomato

More than 300 viruses are able to infect tomato worldwide, including members of the genera *Cucumovirus* (2) (family *Bromoviridae*), *Potyvirus* (12) (family *Potyviridae*), *Polerovirus* (3) (family *Luteoviridae*), *Tymovirus* (2) (family *Tymoviridae*), *Begomovirus* (221) (family *Geminiviridae*), *Crinivirus* (4) (family *Closteroviridae*), *Orthotospovirus* (16) (family *Tospoviridae*) and *Tobamovirus* (6) (family *Virgaviridae*) (GenBank, 2022; Host DATABASE, 2022; Kitajima, 2022; ICTV, 2023). Diseases caused by *Begomovirus* species are among the major biotic problems in Brazil. Diseases induced by begomovirus in tomatoes were first reported after the outbreaks of tomato golden mosaic virus (TGMV) in 1960 (Flores et al., 1960; Matyis et al., 1975). However, it was only in the 1990s that begomoviruses became effectively relevant in the country (Faria et al., 1997; Ribeiro et al., 1998). In this same period, a significant increase *Bemisia tabaci* MEAM 1 (= biotype B) populations (De Barro et al., 2011) was observed, concomitant with the increase in begomoviruses in tomato crops (Ribeiro et al., 1998). *Bemisia tabaci* can cause direct and/or indirect damage to host plants (Hirose et al., 2015). However, the main indirect damage is the ability to transmit viruses from different genera. The whitefly vector can transmit belonging to the genera *Begomovirus*, *Calavirus*, *Crinivirus*, *Ipomovirus*, and *Torradovirus* (Navas-Castillo et al., 2011).

1.4 Family *Geminiviridae*

Geminiviridae family belongs to domain *Monodnaviria*, kingdom *Shotokuvirae*, phylum *Cressdnaviricota*, class *Repensiviricetes* and order *Geplafuvirales*. Currently this family is composed of 14 genera, *Becurtovirus*, *Begomovirus*, *Capulavirus*, *Citlodavirus*, *Curtovirus*, *Eragrovirus*, *Grablovirus*, *Maldovirus*, *Mastrevirus*, *Mulcrilevirus*, *Opunvirus*, *Topilevirus*, *Topocuvirus* and *Turncurtovirus* with of 520 species (Fiallo-Olivé et al., 2021; ICTV 2023) (**Table 1**). *Geminiviridae* is composed of icosahedral twinned particles with size of 22 x 38 nm and circular genome of single strand DNA strand respectively (Fiallo-Olivé et al., 2021). Replication ocurrs by rolling circle mechanism, and mRNA translation for some viral groups of the family uses transcriptional splicing, which consists of removing introns from primary RNA to form a functional mature mRNA (Gutierrez, 1999). Members of *Geminivirdae* are classified as monopartite and bipartite composed of DNA–A or DNA–A and DNA–B, respectivelly (Rojas et al., 2005). They are identified in terms of origin and geographic distribution, New-World viruses derived from the Americas and Old-World viruses derived from Europe, Asia, Africa, and Oceania (Melgarejo et al., 2013). Geminiviruses can be transmitted by treehoppers, leafhoppers, and whiteflies (Fiallo-Olivé et al., 2021). The classification criteria are based upon the host species (monocotyledonous and dicotyledonous) vector species, genome organization, and phylogenetic relationships (ICTV 2023).

Table 1. List of genera of the *Geminiviridae* family considering the number of species by genus, host group, vector, genomic organization, and references.

Genomic organization	**Host**	**Genera (number of species)**	**Vector**	**References**
Monopartite or bipartite	Dicot	*Begomovirus* (445)	*Bemisia tabaci* MEAM 1	Rojas et al., 2005; Fiallo-Olivé et al., 2020.
Monopartite	Monocot	*Mastrevirus* (45)	*Cicadulina mbila*	Muhire et al., 2013; Shepherd et al., 2010; Kvarnheden et al., 2002; Kumar et al., 2014
	Dicot	*Becurtovirus* (3)	*Circulifer haematoceps*	Heydarnejad et al., 2013; Hernández-Zepeda et al., 2013; Razavinejad et al., 2013
		Capulavirus (4)	*Aphis craccivora*	Roumagnac et al., 2015; Ryckebusch et al., 2020; Susi et al., 2019; Bernardo et al., 2013; 2016.
		Citlodavirus (4)	Desconhecido	Loconsole et al., 2012; Zhang et al., 2018; Qiu et al., 2020; Fontenele et al., 2018.
		Curtovirus (3)	*Circulifer tenellus*	Strausbaugh et al., 2008; Hanley-Bowdoin et al., 2013; Gilbertson et al., 2003. Varsani et al., 2014a
		Eragrovirus (1)	Desconhecido	Varsani et al., 2009; Varsani et al., 2014b
		Grablovirus (3)	*Spissistilus festinus*	Krenz et al., 2012; Bahder et al., 2016.
		Maldovirus (3)	Desconhecido	Liang et al., 2015; Al Rwahnih et al., 2017; Claverie et al., 2018.
		Mulcrilevirus (2)	*Tautoneura mori*	Lu et al., 2015; Qui et al., 2020; Lu et al., 2021.
		Opunvirus (1)	*Dactylopius* sp	Fontenele et al., 2020.
		Topilevirus (2)	Desconhecido	Fontenele et al., 2017; Vaghi Medina et al., 2018; Batista et al., 2019.
		Topocuvirus (1)	*Micrutalis maleifera*	Briddon et al., 1996.
		Turncurtovirus (3)	*Circulifer haematoceps*	Razavinejad et al., 2013; Kamali et al., 2016; Hasanvand et al., 2018; Varsani et al., 2014b

1.5 *Begomovirus*

Infection by begomoviruses is a serious threat to the production of several crops of economic importance. These viruses are widely distributed in tropical, subtropical, and temperate regions of the world (Navas-Castillo et al., 2011). *Begomovirus* comprises viruses with either one DNA–A component (monopartite species) or two components (DNA–A and DNA–B in bipartite species) (**Figure 2**). Both components share approximately 200 kb of the sequence called common region (CR), composing the conserved stem-loop with the 5'–TAATATTAC–3' sequence. The genomic organization of ORFs found in monopartites resembles the DNA–A component of bipartite viruses (Hanley-Bowdoin et al., 2000). The DNA–A component has two ORFs in viral sense and five ORFs in the complementary sense (Jeske, 2009; Li et al., 2015). The monopartite viruses and the DNA–A component of bipartite viruses have the AV1/V1 ORF, and monopartite and bipartite begomoviruses from Old World have AV2/V2 (Figure 2). The complementary strand has the ORF Rep (responsible for encoding and forming the replication-associated protein, which initiates replication by cleaving the viral DNA by binding to interacting motifs within the intergenic region and introducing a conserved sequence 5'–TAATATTAC–3' (Laufs et al., 1995, Fontes et al., 1994). Rep also binds to the plant homologue of the protein retinoblastoma, to regulate the cell cycle by altering the environment to favor viral mechanisms. The ORF AC1/ C1 codes for the TrAP which is a viral transcription-activating protein, gene expression in the viral sense of both components, and acts to suppress transcriptional and post-transcriptional gene silencing (Sunter and Bisaro 1992; Bisaro 2006). The ORF AC2/ C2, codes for the replication-enhancing protein vital for the viral replication process (Sunter et al., 1990). The ORFs AC3/C3 and AC4 are extremely important, being related to the symptom expression (Hanley-Bowdoin et al., 2013). The ORF AC5 codes for a protein that is a pathogenicity determinant that is able to suppress antiviral defenses (Li et al., 2015).

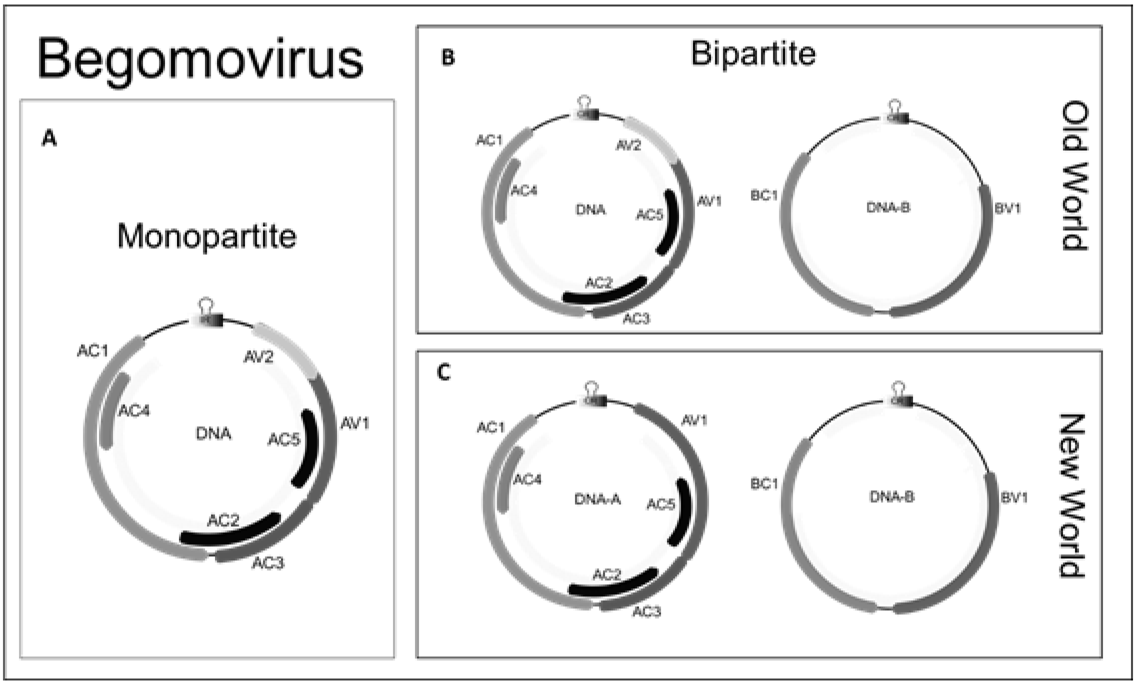


Figure 2 – Genomic organization of *Begomovirus* species: **(A)** Monopartite; **(B)** Bipartite Old World, and **(C)** Bipartite New World. The ORF (Open Reading Frame) V1/AV1, encodes for coat protein (CP); V2/AV2: movement protein (MP); C1/AC1: replication associated protein (Rep); C2/AC2, transativator protein (TrAp); C3/AC3: Enhancer protein (REn) CA/AC4: symptoms expression; C5/AC5: silencing genic supression and pathogenicity; BV1: nuclear shuttle transport (NSP) and BC1: movement protein (MP). The "common region" (RC) that is shared between the two genomic components of the bipartite viruses is displayed as light green boxes within the intergenic region (IR) in monopartite. The position of the stem-loop containing the conserved sequence 5'–TAATATTAC–3' located in the intergenic region (IR) or LIR is shown.

The Rep is composed of subdomains, they are variable amino acids between viruses of different iterons. However, among viruses that have identical iterons, these subdomains are highly conserved. The study of Arguello-Astorga et al. (2001) allowed the grouping of begomoviruses with unique iterons. Another conserved region in Rep has 11 amino acids. This conserved genomic region is related to the host cell cycle, interacting with plant retinoblastoma. Nucleotide insertions or deletions result in altered interaction with retinoblastoma cells. Therefore, this region is considered as extremely relevant for the replication process of begomoviruses in host plants (Arguello-Astorga et al., 2004).

Begomoviruses are efficiently transmitted by polyphagous vector *Bemisia tabaci* Middle East Asian Minor MEAM-1 (= biotype B). This classification is based on the analysis of the nucleotide sequence, mtCOI gene, mitochondrial cytochrome oxidase I, comparing them with consensus regions described for the different species (Dinsdale et al. 2010; De Barro et al., 2011; Polston 2014). This polyphagous vector contributes to genetic variability of begomoviruses under field conditions. Three main mechanisms are responsible to generate genetic variability: pseudo-recombination, recombination and mutation (Roossinck 1997; Seal et al., 2006). **Mutation** is defined by the insertion or deletion or exchange of nucleotides in the gene sequence, this is considered the main mechanism of genetic variability in begomoviruses (Lima et al., 2017). **Recombination** consists of transferring fragments of a sequence from one virus to another, called recombinant. This mechanism is very recurrent in begomoviruses due to mixed infections. Analysis of genetic variability events between tomato yellow vein streek virus and tomato golden vein virus, demonstrated recombination events in the DNA–B, component, among isolates from Brazil, Chile, and Argentina (Reis et al. 2021). Recombination events were also detected in *Macroptilium bright yellow interveinal virus*, involving *bean golden mosaic virus* and *tomato mottle leaf curl virus* (Batista et al., 2022).

1.6 Satellite DNA

Satellite DNAs are consistently associated with Old-World monopartite begomoviruses. However, a satellite DNA group has been reported associated with New-World bipartite begomoviruses (Rojas et al., 2005). Currently, satellite DNAs are classified as alphasatellites (Paprotka et al., 2010; Romay et al., 2010), betasatellites (Mansoor et al., 2003; Saunders et al., 2000) and the begomovirus-dependent Deltasatellites (Ferro et al., 2021).

1.7 Host range of begomoviruses

Begomoviroses can infect a variety of plant species belonging to the dicotyledonous and monocotyledonous (Kil et al., 2021) Many tomato-infesting weed species from families including Amaranthaceae (George et al. 2014), Asteraceae (Tahir et al., 2015), Fabaceae (Batista et al. 2022); Cleomaceae (Fontenele et al. 2017), Cucurbitaceae (Maachi et al., 2022), Euphorbiaceae, Malvaceae, and Solanaceae (Barreto et al., 2013; Duarte et al., 2020; Duarte et al., 2021a) were confirmed as natural hosts of begomoviruses (**Figura 3**). Begomoviruses infect mainly dicotyledonous species. However, monocotyledonous weeds that have been shown to be natural hosts of begomoviruses, such as *Digitaria ciliaris*, *Echinochloa crus-galli*, *Eleusine indica*, *Panicum dichotomiflorum*, and *Setaria faberi* (Kil et al., 2021).

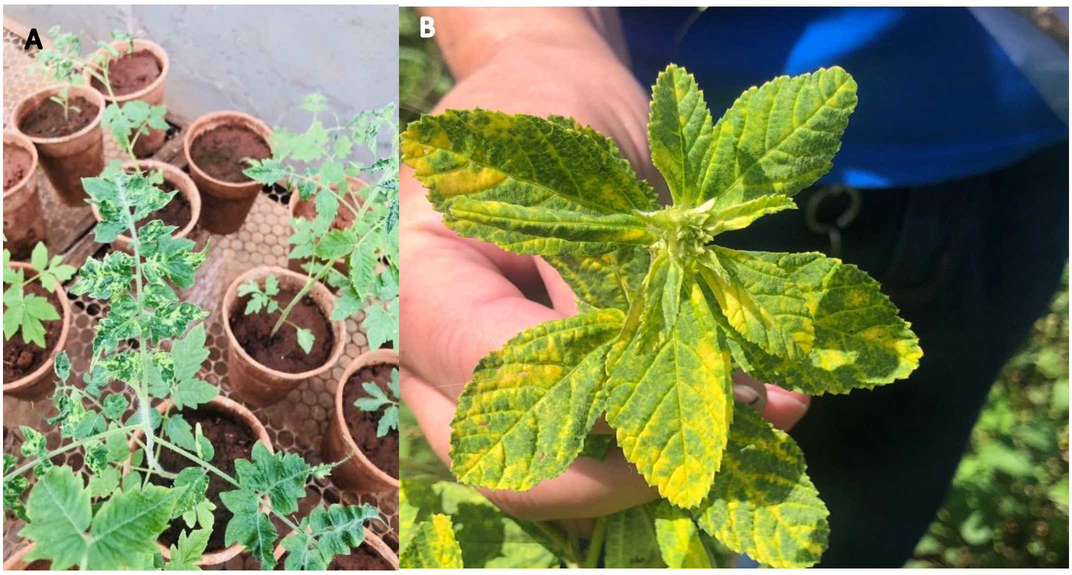


Figure 3 – Plants with typical symptoms of begomoviruses. **A**. Tomato showing symptoms of mosaic and leaf deformation in the apical region. **B**. Malvaceae weeds showing golden mosaic.

1.8 *High-Throughput Sequencing* (HTS)

High-Throughput Sequencing (HTS) is used to produce large amounts of genomic information in a short period of time (Metzker 2005; Posada-Cespedes et al., 2016). Several platforms have been created, such as 454 sequencer, Illumina Genome Analyzer, SOLiD System, Ion Torrent, PacBio, Oxford Nanopore and others (Ambardar et al., 2016; Naito et al., 2019). This set of advanced sequencing technology has been widely used for metagenomics in plant virology. The large-scale genomic evaluations of viruses associated with plants and other organisms using HTS was named as 'virome' (Adams et al., 2009; Barba et al., 2014; Leichtfried et al., 2019; Villamor et al., 2019). In plant virology, HTS has contributed to the identification and characterization of different pre-existing viruses that can affect plants. The application of HTS in plant virology has allowed a more detailed and comprehensive analysis of the pathogens that affect plants, such as the assessment of begomovirus diversity in tomato (Souza et al., 2020; Reis et al., 2020; Duarte et al., 2021b) and verification of host adaptability of viruses (Nehra et al., 2022).

The precise and fast identification of viral pathogens is essential for understanding the disease onset and progression in the host (Lima 2009). In addition, it can be used to analyze the relationships between pathogen and hosts (Souza et al. 2020; Reis et al., 2020). Finally, the use of this tool allows us to catalog plant viruses, their variants, interaction between pathogen and host, so that more efficient strategies can be developed to control viral diseases in plants of economic importance with weeds that play an important role in the monitoring of the viral pathosystem. Hence, the main objective of the present study was to characterize (via HTS) the diversity of the begomoviruses present in tomato-associated weeds classified in the families Malvaceae, Euphorbiaceae, Amaranthaceae, Solanaceae, Asteraceae, Rubiaceae, Fabaceae, Lamiaceae, Cucurbitaceae, Convolvulaceae and Brassicaceae.

REFERENCES

Adams IP, Glover R, Monger W, Mumford R, Jackeviciene E, Navalinskiene M, Samuitiene M, Boonham N. 2009. Next-generation sequencing and metagenomic analysis: A universal diagnostic tool in plant virology. **Molecular Plant Pathology**, 10:537-545.

Al Rwahnih M, Alabi OJ, Westrick NM, Golino, D, Rowhani A. 2017. Description of a novel monopartite geminivirus and its defective subviral genome in grapevine. **Phytopathology**, 107:240-251.

Ambardar S, Gupta R, Trakroo D, Lal R, Vakhlu J. 2016. High Throughput Sequencing: An overview of sequencing chemistry. **Indian Journal Microbiology**, 56:394-404.

Argüello-Astorga GR, Ruiz-Medrano R. 2001. An iteron-related domain is associated to Motif 1 in the replication proteins of geminiviruses: Identification of potential interacting amino acid-base pairs by a comparative approach. **Archives of Virology**, 146:1465-1485.

Arguello-Astorga G, Lopez-Ochoa L, Kong LJ, Orozco BM, Settlage SB, Hanley-Bowdoin L. A novel motif in geminivirus replication proteins interacts with the plant retinoblastoma-related protein. 2004. **Journal of Virology**,78:4817-4826.

Bahder BW, Zalom FG, Jayanth M, Sudarshana MR. 2016. Phylogeny of geminivirus coat protein sequences and digital PCR aid in identifying *Spissistilus festinus* as a vector of grapevine red blotch-associated virus. **Phytopathology**, 106:1223-1230.

Barba M, Czosnek H, Hadidi A. 2014. Historical perspective, development and applications of next-generation sequencing in plant virology. **Viruses**, 6:106-136.

Barreto SS, Hallwass M, Aquino OM, Inoue-Nagata AK. 2013. A study of weeds as potential inoculum sources for a tomato-infecting begomovirus in central Brazil. **Phytopathology**, 103:436-444.

Batista JG, Melo FL, Pereira-Carvalho RC, Alves-Freitas, DMT, Ribeiro SG. 2019. First report of tomato apical leaf curl virus infecting tomato in Brazil. **Plant Disease**, 103:1443.

Batista JG, Nery FMB, Melo FFS, Malheiros MF, Rezende DV, Boiteux LS, Fonseca MEN, de Miranda BEC, Pereira-Carvalho RC. 2022. Complete genome sequence of a novel bipartite begomovirus infecting the legume weed *Macroptilium erythroloma*. **Archives of Virology**, 167:1597-1602.

Bernardo P, Golden M, Akram M, Naimuddin Nadarajan N, Fernandez E, Granier M, Rebelo AG, Peterschmitt M, Martin DP, Roumagnac P. 2013. Identification and characterisation of a highly divergent geminivirus: evolutionary and taxonomic implications. **Virus Research**, 177:35-45.

Bernardo P, Muhire B, François S, Deshoux M, Hartnady P, Farkas K, Kraberger S, Filloux D, Fernandez E, Galzi S, Ferdinand R, Granier M, Marais A, Monge Blasco P, Candresse T, Escriu F, Varsani A, Harkins GW, Martin DP, Roumagnac P. 2016. Molecular characterization and prevalence of two capulaviruses: *Alfalfa leaf curl virus* from France and *Euphorbia caput-medusae latent virus* from South Africa. **Virology**, 493:142-153.

Bezerra-Agasie I, Ferreira G, Ávila A, Inoue-Nagata AK. 2006. First report of *Tomato severe rugose virus* in chili pepper in Brazil. **Plant Disease**, 90:114-114.

Bisaro DM. 2006. Silencing suppression by geminivirus proteins. **Virology**, 344:158-168.

Briddon RW, Bedford ID, Tsai JH, & Markham PG. 1996. Analysis of the nucleotide sequence of the treehopper-transmitted geminivirus, *tomato pseudo-curly top virus*, suggests a recombinant origin. **Virology**, 219:387-394.

Camargo AMMP, Camargo FD, Alves H, Camargo Filho WD. 2006. Desenvolvimento do sistema agroindustrial do tomate. **Informações Econômicas**, 36:53-65.

CBOL Plant Working Group (2009) A DNA barcode for land plants. **Proceedings of the National Academy of Sciences of the United States of America**, 106:12794-12797.

Claverie S, Bernardo P, Kraberger S, Hartnady P, Lefeuvre P, Lett JM, Galzi S, Filloux D, Harkins GW, Varsani A, Martin DP, Roumagnac P. 2018. From spatial metagenomics to molecular characterization of plant viruses: a geminivirus case study. **Advances in Virus Research**, 101:55-83.

Cotrim MADA, Krause-Sakate R, Narita N, Zerbini FM, Pavan MA. 2007. Diversidade genética de begomovírus em cultivos de tomateiro no Centro-Oeste Paulista. **Summa Phytopathologica**, 33:300-303.

Cupertino FP, Costa AS. 1970. Determinação da disseminação do vírus do enrolamento em Batatal para semente pelo uso de plantas indicadoras. **Bragantia**, 29:127-137.

De Barro PJ, Liu SS, Boykin LM, Dinsdale AB. 2011. *Bemisia tabaci*: a statement of species status. **Annual Review of Entomology**, 56:1-19.

Dinsdale A, Cook L, Riginos C, Buckley Y M, De Barro P. 2010. Refined global analysis of *Bemisia tabaci* (Hemiptera: Sternorrhyncha: Aleyrodoidea: Aleyrodidae) mitochondrial cytochrome oxidase 1 to identify species level genetic boundaries. **Annals of the Entomological Society of America**, 103:196-208.

Duarte MF, Fonseca MEN, Boiteux LS, Costa H, Ribeiro BM, Melo FL, Pereira-Carvalho RC. 2020. Identification of *Physalis angulata* (Solanaceae) as a natural alternative weed host of tomato severe rugose virus in Brazil. **Plant Disease**, 104:600.

Duarte MF, Pereira-Carvalho RC, Reis LNA, Rojas MR, Gilbertson RL, Costa H, Boiteux LS, Fonseca MEN. 2021a. Natural infection of tomatoes (*Solanum lycopersicum*) by Euphorbia yellow mosaic virus isolates across four Brazilian states. **Plant Disease**, 105: 518.

Duarte MF, Fonseca MEN, Costa H, Fernandes NAN, Reis A, Boiteux LS, Pereira-Carvalho RC. 2021b. Diversity of tomato-infecting begomoviruses and spatiotemporal dynamics of an endemic viral species of the Brazilian Atlantic rain forest biome. **Virus Genes**, 57: 83-93.

FAOSTAT. 2023. Statistics Division of the Food and Agricultural Organization of the United Nations. http://www.fao.org/faostat/en/#home. Accessed 14 jan 2023.

Faria JC, Souza-Dias JAC, Slack SA, Maxwell DP. 1997. A new geminivirus associated with tomato in the state of São Paulo, Brazil. **Plant Disease**, 81:423-423.

Fernandes-Acioli NAN, Pereira-Carvalho RC, Fontenele RS, Lacorte C, Ribeiro SG, Fonseca MEN, & Boiteux LS. 2011. First report of *Sida micrantha mosaic virus* in *Phaseolus vulgaris* in Brazil. **Plant Disease**, 95:1196-1196.

Ferro CG, Zerbini FM, Navas-Castillo J, Fiallo-Olivé E. 2021. Revealing the complexity of sweepovirus-deltasatellite-plant host interactions: Expanded natural and experimental helper virus range and effect dependence on virus-host combination. **Microorganisms**, 9: 1018.

Fiallo-Olivé E, Lett JM., Martin DP, Roumagnac P, Varsani A, Zerbini FM, Navas-Castillo J. 2021. ICTV virus taxonomy profile: *Geminiviridae* 2021. **The Journal of General Virology**, 102.

Fiallo-Olivé E, Pan LL, Liu SS, Navas-Castillo J. 2020. Transmission of begomoviruses and other whitefly-borne viruses: dependence on the vector species. **Phytopathology**, 110:10-17.

Flores E, Silberschmidt K, Kramer M. 1960. Observations on infectious chlorosis of *Malvaceae* in tomatos in the field. **O Biologico**, 26: 65-69

Fontenele RS, Abreu RA, Lamas NS, Alves-Freitas DMT, Vidal AH, Poppiel RR, Melo FL, Lacorte C, Martin DP, Campos MA, Varsani A, Ribeiro SG. 2018. *Passion fruit chlorotic mottle virus*: Molecular Characterization of a new divergent geminivirus in Brazil. **Viruses**, 10: 169.

Fontenele RS, Lamas NS, Lacorte C, Lacerda ALM, Varsani A, Ribeiro SG. 2017. A novel geminivirus identified in tomato and cleome plants sampled in Brazil. **Virus Research**, 240:175-179.

Fontenele RS, Salywon AM, Majure LC, Cobb IN, Bhaskara A, Avalos-Calleros JA, Argüello-Astorga GR, Schmidlin K, Khalifeh A, Smith K. Schreck J, Lund MC, Köhler M, Wojciechowski MF, Hodgson WC, Puente-Martinez R, Van Doorslaer K, Kumari S, Vernière C, Filloux D, Varsani A. 2020. A novel divergent geminivirus identified in asymptomatic new world cactaceae plants. **Viruses**, 12: 398.

Fontes EP, Eagle PA, Sipe PS, Luckow VA, Hanley-Bowdoin L. 1994. Interaction between a geminivirus replication protein and origin DNA is essential for viral replication. **The Journal of Biological Chemistry**, 269:8459-8465.

George B, Kumar RV, Chakraborty S. 2014. Molecular characterization of *Chilli leaf curl virus* and satellite molecules associated with leaf curl disease of *Amaranthus* spp. **Virus Genes**, 48:397-401.

Giordano LB, Fonseca MEN, Silva JB, Inoue-Nagata AK, Boiteux LS. 2005. Efeito da infecção precoce por *Begomovirus* com genoma bipartido em características de frutos de tomate industrial. **Horticultura Brasileira**, 23, 815-818.

Gutierrez C. 1999. Geminivirus DNA replication. **Cellular and Molecular Life Sciences**, 56:313-329.

Hanley-Bowdoin L, Bejarano ER, Robertson D, Mansoor S. 2013. Geminiviruses: masters at redirecting and reprogramming plant processes. **Nature Reviews Microbiology**, 11: 777-788.

Hanley-Bowdoin L, Settlage SB, Orozco BM, Nagar S, Robertson D. 2000. Geminiviruses: models for plant DNA replication, transcription, and cell cycle regulation. **Critical Reviews in Biochemistry and Molecular Biology**, 35:105-140.

Hasanvand V, Kamali M, Heydarnejad J, Massumi H, Kvarnheden A, Varsani A. 2018. Identification of a new turncurtovirus in the leafhopper *Circulifer haematoceps* and the host plant species *Sesamum indicum*. **Virus Genes**, 54:840-845.

Hasiów-Jaroszewska B, Boezen D, Zwart MP. 2021. Metagenomic studies of viruses in weeds and wild plants: a powerful approach to characterise variable virus communities. **Viruses**, 13:1939.

He YZ, Wang YM, Yin TY, Fiallo-Olivé E, Liu YQ, Hanley-Bowdoin L, Wang XW. 2020. A plant DNA virus replicates in the salivary glands of its insect vector via recruitment of host DNA synthesis machinery. **Proceedings of the National Academy of Sciences of the United States of America**, 117:16928-16937.

Hernández-Zepeda C, Varsani A, Brown JK. 2013. Intergeneric recombination between a new, spinach-infecting curtovirus and a new geminivirus belonging to the genus *Becurtovirus*: first New World exemplar. **Archives of Virology**, 158:2245-2254.

Heydarnejad J, Keyvani N, Razavinejad S, Massumi H, Varsani A. 2013. Fulfilling Koch's postulates for beet curly top Iran virus and proposal for consideration of new genus in the family Geminiviridae. **Archives of Virology**, 158:435-443.

Hirose, E.; Bueno, A.F.; Vieira, S.S.; Gobbi, A.L. 2010. Danos causados por diferentes níveis de infestação de *Bemisia tabaci* Biotipo B (Hemiptera: Aleyrodidae) em soja. Embrapa Soja, 84-86.

IBGE. 2023. Levantamento Sistemático da Produção Agrícola. https://sidra.ibge.gov.br/home/lspa/brasil. Accessed 20 Jan 2023.

ICTV. 2023. International Committee on Taxonomy of Viruses. https://talk.ictvonline.org/. Accessed on 25 Feb, 2023.

Inoue-Nagata AK, Lima MF, Gilbertson RL. 2016. A review of geminivirus (begomovirus) diseases in vegetables and other crops in Brazil: current status and approaches for management. **Horticultura Brasileira**, 34:8-18.

Jeske H. 2009. Geminiviruses. Current Topics. **Microbiology and Immunology**, 331:185-226.

Jones RAC. 2014. Plant virus ecology and epidemiology: Historical perspectives, recent progress and future prospects. **Annals of Applied Biology**, 164:320-347.

Kamali M, Heydarnejad J, Massumi H, Kvarnheden A, Kraberger S, Varsani A. 2016. Molecular diversity of turncurtoviruses in Iran. **Archives of Virology**, 161:551-561.

Kil EJ, Byun HS, Hwang H, Lee KY, Choi HS, Kim CS, Lee S. 2021. Tomato yellow leaf curl virus infection in a monocotyledonous weed (*Eleusine indica*). **The Plant Pathology Journal**, 37: 641-651.

Krenz B, Thompson JR, Fuchs M, Perry KL. 2012. Complete genome sequence of a new circular DNA virus from grapevine. **Journal of Virology**, 86:7715.

Kumar J, Kumar J, Singh SP, Tuli R. 2014. Association of satellites with a mastrevirus in natural infection: complexity of Wheat dwarf India virus disease. **Journal of Virology**, 88:7093-7104.

Kvarnheden A, Lindblad M, Lindsten K, Valkonen JP. 2002. Genetic diversity of wheat dwarf virus. **Archives of Virology**, 147:205-216.

Laufs J, Traut W, Heyraud F, Matzeit V, Rogers SG, Schell J, Gronenborn B. 1995. In vitro cleavage and joining at the viral origin of replication by the replication initiator protein of tomato yellow leaf curl virus. **Proceedings of the National Academy of Sciences of the United States of America,** 92:3879-3883.

Leichtfried T, Dobrovolny S, Reisenzein H, Steinkellner S, Gottsberger RA. 2019. Apple chlorotic fruit spot viroid: A putative new pathogenic viroid on apple characterized by next-generation sequencing. **Archives of Virology,** 164:3137-3140.

Li F, Xu X, Huang C, Gu Z, Cao L, Hu T, Ding M, Li Z, Zhou, X. 2015. The AC5 protein encoded by Mungbean yellow mosaic India virus is a pathogenicity determinant that suppresses RNA silencing-based antiviral defenses. **New Phytologist**, 208:555-569.

Liang P, Navarro B, Zhang Z, Wang H, Lu M, Xiao H, Wu Q, Zhou X, Di Serio F, Li S. 2015. Identification and characterization of a novel geminivirus with a monopartite genome infecting apple trees. **The Journal of General Virology**, 96:2411-2420.

Lima ATM, Silva JCF, Silva FN, Castillo-Urquiza GP, Silva FF, Seah YM, Mizubuti ESG, Duffy S, Zerbini FM. 2017. The diversification of begomovirus populations is predominantly driven by mutational dynamics. **Virus Evolution**, 3:1-14.

Lima MFL. 2009. Detecção e controle de viroses em videira. Petrolina: Embrapa Semiárido.9 p

Loconsole G, Saldarelli P, Doddapaneni H, Savino V, Martelli G P, Saponari M. 2012. Identification of a single-stranded DNA virus associated with citrus chlorotic dwarf disease, a new member in the family *Geminiviridae*. **Virology**, 432:162-172.

Lopes CA, Ávila AC. 2005. Doenças do Tomateiro, Embrapa Hortaliças. Brasília–DF.

Lopes CA, Reis A. 2011. Doenças do tomateiro cultivado em ambiente protegido. In. Corcórdia: Embrapa Hortaliças, Brasília–DF, pp. **Circular técnica**, 100.

Lopes CA, Reis A. 2017. Doenças do tomateiro cultivado em ambiente protegido. Embrapa Hortaliças, Brasília–DF, Circular Técnica 100:26.

Lu QY, Wu ZJ, Xia ZS, Xie LH. 2015. Complete genome sequence of a novel monopartite geminivirus identified in mulberry (*Morus alba* L.). **Archives of Virology**, 160: 2135-2138.

Lu QY, Ma Y, Smith WK, Yu J, Cheng YY, Zhang P, Han TT. 2022. The Identification of *Tautoneura mori* as the vector of Mulberry crinkle leaf virus and the infectivity of infectious clones in mulberry. **Phytopathology**,112: 435-440.

Maachi A, Donaire L, Hernando Y, Aranda MA. 2022. Genetic differentiation and migration fluxes of viruses from melon crops and crop edge weeds. **Journal of Virology**, 96: e0042122.

Macedo MA and Barreto SS and Costa TM, Maliano MR, Rojas MR, Gilbertson RL, Inoue-Nagata AK. 2017a. First report of common beans as a nonsymptomatic host of tomato severe rugose virus in Brazil. **Plant Disease**, 101:261.

Macedo MA, Barreto SS, Costa TM, Maliano MR, Rojas MR, Gilbertson RL, Inoue-Nagata AK. 2017b. First report of tomato severe rugose virus, a tomato-infecting begomovirus, in soybean plants in Brazil. **Plant Disease**, 101:1959-1959.

Mansoor S, Briddon RW, Zafar Y, & Stanley J. 2003. Geminivirus disease complexes: an emerging threat. **Trends in Plant Science**, 8:128-134.

Matyis JC, Silva DM, R. OAA, S. CA. 1975. Purificação e morfologia do vírus do mosaico dourado do tomateiro. **Summa Phytopathologyca**, 1:267-274.

Melgarejo TA, Kon T, Rojas MR, Paz-Carrasco L, Zerbini, FM, & Gilbertson RL. 2013. Characterization of a new world monopartite begomovirus causing leaf curl disease of tomato in Ecuador and Peru reveals a new direction in geminivirus evolution. **Journal of Virology**, 87:5397-5413.

Metzker ML. 2005. Emerging technologies in DNA sequencing. **Genome Research**, 15:1767-1776.

Moura MF, Ruschel RG, Gotardi GA, Watanabe LFM, Rêgo CM, Inoue-Nagata AK, Pavan MA, Rezende JAM, Krause-Sakate R. 2018. First report of tomato severe rugose virus in eggplant. **Journal of Plant Pathology**, 100:599-599.

Muhire B, Martin DP, Brown JK, Navas-Castillo J, Moriones E, Zerbini FM, Rivera-Bustamante R, Malathi VG, Briddon RW, Varsani A. 2013. A genome-wide pairwise-identity-based proposal for the classification of viruses in the genus *Mastrevirus* (family *Geminiviridae*). **Archives of Virology**, 158:1411-1424.

Naito FY, Melo FL, Fonseca MEN, Santos CA, Chanes CR, Ribeiro BM, Gilbertson RL, Boiteux LS, Pereira-Carvalho RC. 2019. Nanopore sequencing of a novel bipartite New World begomovirus infecting cowpea. **Archives of Virology**, 164, 1907-1910.

Navas-Castillo J, Fiallo-Olivé E, Sánchez-Campos S. 2011. Emerging virus diseases transmitted by whiteflies. **Annual Review of Phytopathology**, 49:219-248.

Nehra C, Verma RK, Petrov NM, Stoyanova MI, Sharma P, Gaur RK. 2022. Computational analysis for plant virus analysis using next-generation sequencing. **Bioinformatics in Agriculture**, 24:383-398.

Ong SN Taheri S, Othman RY, Teo CH. 2020. Viral disease of tomato crops (*Solanum lycopersicum* L.): An overview. **Journal of Plant Diseases and Protection**, 127: 725-739.

Paprotka T, Metzler V, Jeske H. 2010. The first DNA 1-like alpha satellites in association with New World begomoviruses in natural infections. **Virology**, 404:148-157.

Pereira-Carvalho RC, Tobar LLM, Dianese EC, Fonseca MEN, Boiteux LS. 2014. Melhoramento genético do tomateiro para resistência a doenças de etiologia viral: Avanços e perspectivas. **Revisao Anual de Patologia de Plantas,** 22: 280-361.

Pereira-Silva J, Boiteux LS, Fonseca MEN, Reis LNA, Souza AS, Nery FMB, Madeira NR, Pereira-Carvalho RC. 2022. Novel natural hosts of tomato severe rugose virus (ToSRV) in the Fabaceae, Solanaceae, and Oxalidaceae families. **Journal of Plant Diseases and Protection**, 129: 425-431.

Pinheiro JB, Boiteux LS, Pereira RB, Almeida MRA, Carneiro RMG. 2014. Identificação de espécies de *Meloidogyne* em tomateiro no Brasil. **Embrapa Hortaliças-Boletim de Pesquisa e Desenvolvimento** (INFOTECA-E)

Polston JE, De Barro P, Boykin LM. 2014. Transmission specificities of plant viruses with the newly identified species of the *Bemisia tabaci* species complex. **Pest Management Science**, 70: 1547-1552.

Posada-Cespedes S, Seifert D, Beerenwinkel N. 2017. Recent advances in inferring viral diversity from high-throughput sequencing data. **Virus Research**, 239:17-32.

Qiu Y, Zhang S, Yu H, Xuan Z, Yang L, Zhan B, Murilo Zerbini F, Cao M. 2020. Identification and characterization of two novel geminiviruses associated with paper Mulberry (*Broussonetia papyrifera*) Leaf Curl Disease. **Plant Disease**, 104:3010-3018.

Razavinejad S, Heydarnejad J, Kamali M, Massumi H, Kraberger S, Varsani A. 2013. Genetic diversity and host range studies of turnip curly top virus. **Virus Genes**, 46: 345-353.

Reis LNA, Fonseca MEN, Ribeiro SG, Naito FYB, Boiteux LS, Pereira-Carvalho RC. 2020. Metagenomics of neotropical single-stranded DNA viruses in tomato cultivars with and without the *Ty*-1 gene. **Viruses,** 12(8), 819.

Reis LNA, Boiteux LS, Fonseca MEN, Pereira-Carvalho RC. 2021. Tomato yellow vein streak virus and tomato golden vein virus: A reappraisal of the classification status of two South American *Begomovirus* species based upon genome-wide pairwise identity of multiple isolates. **Virus Genes**, 57:127-131.

Ribeiro SG, de Ávila AC, Bezerra IC, Fernandes JJ, Faria JC, Lima MF, Gilbertson RL, Maciel-Zambolim E, Zerbini FM. 1998. Widespread occurrence of tomato geminiviruses in Brazil, associated with the new biotype of the whitefly vector. **Plant Disease**, 82:830-830.

Richet C, Kraberger S, Filloux D, Bernardo P, Harkins GW, Martin DP, Roumagnac P, Varsani A. 2019. Novel circular DNA viruses associated with *Apiaceae* and *Poaceae* from South Africa and New Zealand. **Archives of Virology**, 164:237-242.

Rojas MR, Gilbertson RL, Maxwell DP (1993) Use of degenerate primers in the polymerase chain reaction to detect whitefly-transmitted geminivíruses. **Plant Disease**, 77:340–347

Rojas MR, Hagen C, Lucas WJ, Gilbertson RL (2005) Exploiting chinks in the plant's armor: evolution and emergence of geminiviruses. **Annual Review of Phytopathology**, 43:361-394.

Romay G, Chirinos D, Geraud-Pouey F, Desbiez C. 2010. Association of an atypical alphasatellite with a bipartite New World begomovirus. **Archives of Virology**, 155:1843-1847.

Roossinck MJ. 1997. Mechanisms of plant virus evolution. Annual Review of **Phytopathology**, 35:191-209.

Roumagnac P, Granier M, Bernardo P, Deshoux M, Ferdinand R, Galzi S, Fernandez E, Julian C, Abt I, Filloux D, Mesléard F, Varsani A, Blanc S, Martin D P, Peterschmitt M. 2015. Alfalfa leaf curl virus: an aphid-transmitted geminivirus. **Journal of Virology**, 89: 9683-9688.

Ryckebusch F, Sauvion N, Granier M, Roumagnac P, Peterschmitt M. 2020. Alfalfa leaf curl virus is transmitted by aphis craccivora in a highly specific circulative manner. **Virology**, 546: 98-108.

Saunders K, Bedford ID, Briddon RW, Markham PG, Wong SM, Stanley J. 2000. A unique virus complex causes Ageratum yellow vein disease. **Proceedings of the National Academy of Sciences of the United States of America**, 97:6890-6895.

Seal S, VandenBosch F, Jeger M. 2006. Factors influencing begomovirus evolution and their increasing global significance: implications for sustainable control. **Critical Reviews in Plant Sciences**, 25:23-46.

Shepherd DN, Martin DP, Van Der Walt E, Dent K, Varsani A, Rybicki EP. 2010. Maize streak virus: An old and complex ‘emerging’ pathogen. **Molecular Plant Pathology**, 11:1-12.

Soto MJ, Gilbertson RL. 2003. Distribution and rate of movement of the curtovirus beet mild curly top virus (Family *Geminiviridae*) in the Beet Leafhopper. **Phytopathology**, 93:478-484.

Souza TA, Silva JMF, Nagata T, Martins TP, Nakasu EYT, Inoue-Nagata AK. 2020. A temporal diversity analysis of Brazilian begomoviruses in tomato reveals a decrease in species richness between 2003 and 2016. **Frontiers in Plant Science**, 11:1-13.

Strausbaugh CA, Wintermantel WM, Gillen AM, Eujayl IA. 2008. Curly top survey in the Western United States. **Phytopathology**, 98:1212-1217.

Sunter G, Bisaro DM. 1992. Transactivation of geminivirus AR1 and BR1 gene expression by the viral AL2 gene product occurs at the level of transcription. **The Plant Cell**, 4:1321-1331.

Sunter G, Hartitz MD, Hormuzdi SG, Brough CL, Bisaro DM. 1990. Genetic analysis of tomato golden mosaic virus: ORF AL2 is required for coat protein accumulation while ORF AL3 is necessary for efficient DNA replication. **Virology**, 179:69-77.

Susi H, Filloux D, Frilander MJ, Roumagnac P, Laine AL. 2019. Diverse and variable virus communities in wild plant populations revealed by metagenomic tools. **PeerJ**, 7: e6140.

Tahir M, Amin I, Haider MS, Mansoor S, Briddon RW. Ageratum enation virus-a begomovirus of weeds with the potential to infect crops. 2015. **Viruses**,7: 647-665.

Vaghi Medina CG, Teppa E, Bornancini VA, Flores CR, Marino-Buslje C, López Lambertini PM. 2018. Tomato apical leaf curl virus: A novel, monopartite geminivirus detected in tomatoes in Argentina. **Frontiers in Microbiology**, 8:2665.

Varsani A, Martin DP, Navas-Castillo J, Moriones E, Hernández-Zepeda C, Idris A, Murilo Zerbini F, Brown JK. 2014a. Revisiting the classification of curtoviruses based on genome-wide pairwise identity. **Archives of Virology**, 159:1873-1882.

Varsani A, Navas-Castillo J, Moriones E, Hernández-Zepeda C, Idris A, Brown JK, Murilo Zerbini F, Martin DP. 2014b. Establishment of three new genera in the family *Geminiviridae*: *Becurtovirus*, *Eragrovirus* and *Turncurtovirus*. **Archives of Virology**, 159:2193-2203.

Varsani A, Shepherd DN, Dent K, Monjane AL, Rybicki EP, Martin DP. 2009. A highly divergent South African geminivirus species illuminates the ancient evolutionary history of this family. **Virology Journal**, 6:1-12.

Villamor D, Ho T, Al Rwahnih M, Martin R, Tzanetakis I. 2019. High throughput sequencing for plant virus detection and discovery. **Phytopathology**, 109:716- 725.

Zhang S, Shen P, Li M, Tian X, Zhou C, Cao M. 2018. Discovery of a novel geminivirus associated with camellia chlorotic dwarf disease. **Archives of Virology**, 163:1709-1712.

CHAPTER 2

Metagenomic analyses and diversity of single-stranded DNA viruses infecting Neotropical weeds collected within and in the edge of tomato fields.

.

To be submitted to Virology Journal

Metagenomic analyses and diversity of single-stranded DNA viruses infecting Neotropical weeds collected within and in the edge of tomato fields.

Eduardo Soares da Silva Lima[1], Luciane de Nazaré Almeida dos Reis[1], Felipe Fochat Silva Melo[1], Leonardo S. Boiteux[2], Maria Esther de Noronha Fonseca[2], Rita de Cássia Pereira-Carvalho[1*]

[1]Departamento de Fitopatologia, Universidade de Brasília (UnB), Brasília–DF, Brazil.

[2]National Center for Vegetable Crops Research (CNPH), Embrapa Vegetable Crops (Embrapa Hortaliças), Brasília–DF, Brazil.

*Correspondence author: rcpcarvalho@unb.br

Abstract

Background

Begomoviruses can induce significant yield losses ranging from 40–100% in tomato (*Solanum lycopersicum* L.) in Neotropical areas. Tomato-associated weeds may act as begomovirus reservoirs and play essential roles in viral biodiversity. Howver, viromes of tomato weeds are yet scarce. Herein, High-Throughput Sequencing (HTS) was employed to generate a comprehensive ssDNA virome using symptomatic samples of neotropical weeds occurring in association with tomato fields in the five Brazilian macro geographical regions. The main objective was to catalog the diversity and to verify the role of weeds from 11 botanic families as potential reservoirs of tomato-infecting ssDNA viruses.

Methods

Metagenomic analyses of circular, single-stranded DNA (ssDNA) viruses (family *Geminiviridae*) of neotropical weeds associated with tomato (*Solanum lycopersicum* L.) fields were carried out using HTS. Ninety-one weed samples exhibiting typical geminivirus-like symptoms were evaluated in species of the families Amaranthaceae (04), Asteraceae (03), Brassicaceae (01), Cleomaceae (01), Concolvulaceae (01), Cucurbitaceae (01), Euphorbiaceae (03), Fabaceae (02), Lamiaceae (01), Malvaceae (71), and Solanaceae (03). After DNA extraction of individual samples, Rolling Circle Amplification (RCA) was performed for enrichment of circular genomes of viral origin. Individual RCA samples were mixed into a single pool that was sequenced at an Illumina NovaSeq6000 platform.

Results

A total of 16079516 reads and 38901 contigs were obtained, with 100 corresponding to viral genome sequences with identity to members of three genera (*Begomovirus*, *Topilevirus*, and *Mulcrilevirus*). Sixteen of the 94 contigs with identity to the genus *Begomovirus* corresponded to 14 putative new viral species. The remaining 86 contigs displayed identity to previously characterized begomoviruses. Four contigs (C393, C708, C791, and C1259), with 100% identity among them, showed 97.91% identity with tomato apical leaf curl virus – ToALCV (*Topilevirus*). One contig (C097) showed 92.25% identity with mulberry crinckle associated virus – MCaV (*Mulcrilevirus*). In addition, one contig (named as 151) showed 96.12% identity with Euphorbia yellow mosaic alpha satellite – EYMA (genus *Clecrusatellite*).

Conclusions

The present work confirms the extraordinary potential of neotropical weeds as reservoirs of circular ssDNA viruses and satellites that might represent potential pathogens of tomato and other hosts of economic importance.

Keywords: *Begomovirus*, *Topilevirus*, *Mulcrilvirus*, *Clecrusatellite*, ssDNA, *High-Throughput Sequencing*

BACKGROUND

The *Geminiviridae* family comprises non-enveloped single-stranded DNA viruses with genomes ranging from 2.5 to 5.2 kb, which are encapsidated into incomplete icosahedral twinned particles with a size of 22 *x* 38 nm [1]. These viruses exhibit a wide host range, including crops of economic importance, as well as ornamentals and weed plants. *Begomovirus* is among the 14 genera of the family with largest number of members in the family *Geminiviridae* [2]. Begomoviruses are characterized by either monopartite (only DNA–A) or bipartite (DNA–A and DNA–B) genomes. The Open Reading Frames (ORFs) found in the monopartite genomes resemble the ones found in the DNA–A of the bipartite species. Both monopartite and bipartite DNA–A genomes can contain up to seven ORFs. The ORF V1/AV1 codes for the coat protein (CP), which is an essential protein for the interaction of the virus with the vector and for efficient transmission [3, 4, 5, 6, 7]. *Begomovirus* of the Old World presented the ORF V2/AV2 involved in viral movement in the host cell, allowing the virus to reach different cells as well as the phloem [8]. In the complementary sense of DNA–A, five ORFs are found. The ORF C1/AC1 codifying the protein associated with replication (Rep), able to recognize and bind to viral DNA, and prepare DNA for replication, and synthesis of new DNA molecules and catalyzing the integration of viral DNA into the host genome [9, 10]. The ORF C2/AC2 encodes a transcription activator protein (TrAp), which is a multifunctional protein playing important roles in modulation of viral gene expression, regulation of host plant immune response and virus virulence [11, 12]. The ORF C3/AC3 encode a replication enhancer protein (Ren) [13]. The ORF C4/AC4 encode a suppressor of the host plant defense response, important in the process of expression of symptoms and alteration of the cell life cycle [14, 15], and the ORF C5/AC5 is able to inhibit the transcription of plant defense genes and prevents the plant's immune response to the virus [16]. In the viral sense of the DNA–B component, the ORF BV1 codes for the nuclear shuttle protein (NSP), important in nuclear transport of DNA [17]. The ORF BC1 encodes the movement protein (MP), which plays the role of virus movement between cells and tissues of the plant host [17, 18]. Begomoviruses are transmitted by members of the *Bemisia tabaci* cryptic species complex (Hemiptera: Sternorrhyncha: Aleyrodoidea: Aleyrodidae), which comprises more than 44 morphologically identical species [19]. The definition of species in this complex is done by comparing the levels of sequence identity of mitochondrial cytochrome oxidase I (COI) across insect populations [20]. The insect acquires the virus in the infected plant, which travels through the digestive tract, intestine, through the hemolymph where it gains access to the salivary glands [21].

Begomoviruses can induce significant yield losses ranging from 40–100% in tomato (*Solanum lycopersicum* L.) under Brazilian conditions [22]. Nationwide surveys provide relevant information for anticipatory breeding programs, monitoring the present and potential variability of geminiviruses in major producing areas infecting either tomato or associated weeds [23, 24, 25, 26, 27, 28, 29]. In this context, the use of High-Throughput Sequencing (HTS) platforms has been a fundamental tool for the study of viral diversity as well as the discovery of new viruses in different hosts of economic importance and their associated weeds. These data also provide an understanding of the pathosystem, as well as the necessary information for the construction of more effective control methodologies.

According to Rodríguez-Negrete et al. (2019), HTS can provide information on the diversity of begomovirus hosts [30]. The data obtained reported that 60% belong to the group of invasive plants in the North Pacific of Mexico, corroborating that these serve as a source of inoculum for plants of agricultural importance. In studies by Rivarez et al. (2023) three new viruses classified in families *Potyviridae*, *Rhabdoviridae*, and *Tombusviridae* were also discovered from hosts of the Solanaceae and Asteraceae [31]. Beside this, new genera of viruses have been discovered using HTS, including *Topilevirus* [32] and *Mucrilevirus* [33].

Such findings support the notion that weeds act as begomovirus reservoirs and play essential roles in begomoviruses biodiversity. Therefore, the control of their populations close to commercial crops should be considered to avoid the harmful effects of these plant pathogens and thus increase agricultural efficiency, ensure food and nutritional security. In this work, a metagenomic analysis of viruses found in weeds associated with tomato cultivation was carried out using HTS, located in the five Brazilian macro geographical regions. Here, HTS was employed to generate a comprehensive ssDNA virome using symptomatic samples of neotropical weeds occurring in association with tomato fields. The main objective was to catalog the diversity and to verify the role of weeds from 11 botanic families as potential reservoirs of tomato-infecting ssDNA viruses.

METHODS

Collection and maintenance of weed samples – Ninety-one symptomatic leaf samples classified in 11 botanical families: Amaranthaceae, Asteraceae, Brassicaceae, Cleomaceae, Concovulaceae, Cucurbitaceae, Euphorbiaceae, Fabaceae, Lamiaceae, Malvaceae, and Solanaceae, were collected close to tomato crops between 2003 and 2022 in all macroregions of Brazil (**Supplementary Table 1**). These samples were selected using as a criteria location and year, and the set of samples belonging to the collection of 1400 isolates of the Breeding Laboratory at CNPH, Brasília–DF (Federal District).

Total DNA extraction, PCR with universal primers, and amplification of circular ssDNA viruses via RCA – The genomic DNA was extracted using a protocol with 2X CTAB and organic solvents [34]. The total DNA was used as template for the Rolling Circle Amplification (RCA) procedure [35]. Initially PCR was performed to confirm the presence of begomoviruses, based on the partial sequence of ≅ 1200 nucleotides obtained by universal primers 'PAL1v1978' (5'–GCA TCT GCA GGC CCA CAT YGT CTT YCC NGT–3') and 'PAR1c496' (5'–CAT GCT GCA GTA CAT YGG CCT YTT DAC CC–3') [36]. The protocol for preparing the reaction mix was for a total volume of 12.5 µL with 9 µL Milli-Q water, 1.25 µL of 10X *Taq* Polymerase enzyme buffer, 0.4 µL magnesium chloride 50 mM ($MgCl_2$), 0.25 µL of dNTPs 10 mM, 0.25 µL of each primer 10 mM, 0.1 µL *Taq* DNA polymerase enzyme, and 1.0 µL of DNA. The PCR mix was submitted to the thermal cycler with the following conditions, 94ºC for 3 minutes, followed by 35 cycles, each one consisting of: 94ºC for 30 seconds for denaturation, 50ºC for 1 minute for annealing and 72ºC for 1 minute for extension, and a final extension step at 72ºC for 7 minutes. The amplicons were subjected to electrophoresis in 1% agarose gel, stained with ethidium bromide and visualized under ultraviolet light.

HTS analysis and phylogenetic analyses – The RCA products of all samples were grouped into a single pool (**Supplementary Table 1**). After pool preparation, the samples were submitted to HTS sequencing on the Illumina NovaSeq-6000 platform. The raw reads were quality trimmed and assembled using CLC Genomics workbench v.8.0 (Qiagen). The assembled contigs were analyzed by BLASTn algorithm against the viral RefSeq database at the GenBank (https://www.ncbi.nlm.nih.gov/). The contigs related to plant viruses were selected. Based on these results, the complete genome was assembled using the Geneious® 11.1.5 program [37]. The reads obtained were mapped to the contig of a potential virus to obtain the final genome. The genomes of the individual contigs were extended using the Geneious® program and the Map to reference tool (parameter 90 to 99% of minimum overlap identity) with mapping in the reads file provided by the HTS. All contigs were subjected to comparisons with viral sequences present in the GenBank using BLASTn algorithm. MUSCLE alignments were performed in the Geneious program for annotation of ORFs based on the reference genome. The Geneious® R 11.1.5 program was also used to assemble the viral genome, annotate, and align the assembled sequences. Sequences were aligned in Geneious® 11.1.5 using MUSCLE plugin with reference genomes available within the GenBank database. Maximum likelihood phylogenetic analysis of both nucleotide and amino-acid sequences were performed with Geneious® plugin PhyML [38] using GTR model [39] and 1,000 bootstrap replications.

RESULTS

High-throughput sequencing provided the following raw data: 16,079,516 reads and 38,901 contigs of which 100 corresponded to viral genome sequences. Of these, 94 contigs corresponded to viruses classified in *Begomovirus*, four contigs in *Topilevirus*, one contig for *Mulcrilevirus* and, one contig for *Clecrusatellite*. Among the 94 contigs of begomoviruses, 48 contigs corresponded to DNA–A (**Table 1)** while 46 contigs to DNA–B (**Table 2**). Of these 48 contigs, 20 had identity below 91%, indicating that they may represent novel species. However, six of these 20 contigs showed high identity among themselves. Contigs C150 and C16 showed 95.47% identity between them. Contigs C490 and C98 showed 93% identity, whereas the contigs C43 and C46 showed 97.25% identity between them. The other 28 contigs showed identity with viruses reported in the families Malvaceae (15), Euphorbiaceae (4), Fabaceae (2), Solanaceae (6), and Cleomaceae (1).

For the Malvaceae family, 15 contigs were recovered and they corresponded to Abutilon mosaic virus (AbMV) and Sida micrantha mosaic virus (SiMMV). For the Euphorbiaceae family, all four contigs (named as C40, C48, C153 and C3087) recovered corresponded to Euphorbia yellow mosaic virus (EuYMV). For the Solanaceae family, the six contigs recovered corresponded to three viruses: tomato mild mosaic virus (ToMMV), tomato severe rugose virus (ToSRV), and tomato yellow net virus (ToYNV). For the Fabaceae family, the two recovered contigs corresponded to bean golden mosaic virus (BGMV). For the Cleomaceae family, a contig was recovered corresponding to Cleome golden mosaic virus (CleGMV). In addition to the begomoviruses, viruses classified as *Mulcrilevirus*, *Topilevirus*, and *Clecrusatellite* were found. For *Mulcrilevirus*, the C97 contig showed 92.25% identity with the mulberry crinkle-associated virus (MCAV). The contig C151 was classified in the *Alphasatellitidae* family and showed 96.12% identity with Euphorbia yellow mosaic alphasatellite (EYMA). For the genus *Topilevirus*, four contigs (named as C791, C393, C708 and C1259) recovered corresponded to tomato apical leaf curl virus (ToALCV) (**Table 3**). The analysis using SDT performed with sequences of already known species and those obtained in this work, confirmed the detection of 15 new species showing an identity of 70% to 90% between them (**Figure 1**). Phylogenetic analyzes indicated that most of these new species were close to begomoviruses species originally detected in weeds, with the exception of contig 88, which was closer to the tomato ToSRV species (**Figure 2**).

CONCLUSIONS

HTS sequencing makes it possible to obtain an immense amount of viral sequence data, which can generate up to 6000 gigabytes of data, with up to 20 billion reads per cycle and 75% accuracy [40]. Herein, the sequencing and assembling the contigs from this sample pool provided data for the recovery of 100 viral genomes. From these assembled sequences, viruses belonging to four genera were recovered: *Mulcrilevirus*, *Clecrusatellite*, *Topilevirus*, and *Begomovirus*.

Contig 97 showed 92.25% identity with the mulberry crinkle-associated virus (MCAV) of the genus *Mulcrilevirus* (**Table 3**). Viruses of this genus are characterized by single-stranded DNA genomes with genomic organization containing seven ORFs, V1, V2, V3, V4, C1 and C2 and C1 [41]. However, it had not yet been reported in the American continent. Being described in China in the year 2015 Mulberry crinkle leaf virus [33, 42] and Paper mulberry leaf curl virus 1 [41] both in mulberry.

In these results, HTS recovered EYMA with 1337 kb belonging to the genus *Clecrusatellite* (**Table 3**). Possibly this alphasatellite infected the host through a helper begomovirus, knowing that it lacks essential proteins for recognizing the vector and enabling its transmission [43]. According to Mar et al. (2017), coinfection of EYMA with EYMV in *N. benthamiana* and *E. heterophylla* favored the severity of symptoms [44]. As stated by Nawaz-ul-Rehman et al. (2010), the replication protein may have acted as a suppressor of gene silencing, thus increasing the severity of the disease [45]. In 2010 and 2014, the natural presence of alphasatellite in weeds associated with bipartite begomoviruses was reported [46, 47]. The EuYMV virus was also identified in four contigs of this work. This virus was detected in *E. heterophylla* (belonging to the Euphorbiaceae family), as well as in *S. lycopersicum* [24, 44].

For the genus *Topilevirus*, four contigs were recovered that showed high identity with the ToALCV virus (**Table 3**). This virus has *S. lycopersicum* as a typical host, however this group of samples does not have tomatoes, they only have two hosts of the Solanaceae family, which are Muricato and Jurubeba, taxonomically related species. In Brasília, in 2016, ToALCV was detected and characterized molecularly and biologically in tomato, being the first report of this virus in Brazil [32].

The data provided enabled the assembly of 94 sequences related to viruses of the *Begomovirus* genus, 48 of which were DNA–A (**Table 1**) and 46 were DNA–B sequences (**Table 2**). The existence of more DNA–A in relation to DNA-B may indicate that more than one A-component is using the same DNA–B for viral movement. The existence of mixed infections can contribute to the process of pseudo-recombination, which is the mechanism of exchange of genomic components between different viral species [48, 49]. Furthermore, 20 of these 48 contigs referring to DNA–A showed low identity, which corresponded to 15 new species. For begomoviruses, the species demarcation criterion corresponds to the identity of 91% of shared nucleotides in the DNA in monopartite or DNA–A in bipartite [50].

Several factors contribute to generate genetic diversity leading to the potential emergence of new geminiviruses. Studies have already reported that geminiviruses in have a high rate of nucleotide substitution [51]. In addition, recombination is recurrent genetic mechanism in this group of viruses, contributing to viral speciation [52]. The cosmopolitan distribution of the polyphagous vectors also favors interactions among distinct viruses derived from different hosts in mixed infections in a single host cell, increasing the chances of generating genetic variability and the emergence of new species [53].

Many invasive and introduced plant species belonging to the botanical families Asteraceae, Caparaceae, Euphorbiaceae, Fabaceae, Lamiaceae, Malvaceae, Solanaceae, and Sterculiaceae have been reported as hosts of viruses belonging to the genus *Begomovirus* in the Americas and Caribbean region [54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66].

For the Fabaceae family, two contigs were found that showed high identity with BGMV, this virus is considered one of the most important today, as it can cause losses of up to 100% of bean production. The main host of BGMV is *Phaseolus vulgaris*, for which no effective natural resistance gene has been detected [67]. However, BGMV has been reported infecting other hosts such as, *Glycine max*, *Macroptilium atropurpureum*, *Canavalia essiformis*, *Phaseolus lunatus*, and *Lathyrus sativus* [68, 69]. For the Cleomaceae family, the Cleome golden mosaic virus was detected, which has three sequences of this virus obtained from *Cleome* species deposited at the GenBank (https://www.ncbi.nlm.nih.gov/).

For the Solanaceae family, six contigs related to three viruses were detected. The tomato yellow net virus (ToYNV), tomato mild mosaic virus (ToMIMV) and tomato severe rugose virus (ToSRV). The ToYNV virus was found only in tomato, being a species not yet accepted by the ICTV. In this work, the report of this species in samples from other botanical families will be a novelty, knowing that obtaining this information is very important for further research, characterization and understanding of this virus [27]. The ToMIMV species has already been reported in two hosts, S. *lycopersicum* and *Sida urens*, currently presenting 15 sequences deposited in the database [70]. Only one contig was related to the ToSRV virus, which is of great importance for tomato, having been found in virtually all regions of the country. Recently, ToSRV was detected in *Physalis angulata* that it can function as a natural source of inoculum within tomato production fields. The control of invasive plants should be part of the integrated management of this virus species in tomato crops [24].

For the Malvaceae, 15 contigs showed high identity for SiMMV and AbMV. The SiMMV was initially reported in a weed and described in *P. vulgaris* in Brazil [71]. AbMV was identified as the agent responsible for the common mosaic in cotton cultivation in the state of Goiás [72]. The Malavaceae family has the largest diversity of natural host species of begomoviruses, among these species are plants of economic importance such as *Abutilon* and *Sida* species [73].

The tomato is a relevant vegetable worldwide. There are several reports of Neotropical weeds playing a role as reservoirs of tomato yield-reducing circular, single-stranded DNA (ssDNA) viruses (family *Geminiviridae*), especially members of the genus *Begomovirus*. Here, HTS was employed to generate a comprehensive circular ssDNA virome using symptomatic samples of neotropical weeds from 11 botanical families occurring in association with tomato fields in all macro-geographic regions of Brazil. The present work confirms the extraordinary potential of tomato crop-associated weeds as reservoirs of novel as well as previously described circular ssDNA viruses and satellites that might represent potential pathogens of tomatoes and other hosts of economic importance.

Table 1 – Number of contigs, reading coverage, assembled genome size, BLASTn coverage, sequence identity of the assembled virus, E-value, GenBank (GB) accession of begomoviruses discovery by High-Throughput Sequencing (HTS) in a pool of 91 foliar samples of weeds (from 11 botanic families), containing DNA–A segments of distinct *Begomovirus* species.

Contigs	**Coverage reads**	**Size (nucleotides)**	**Coverage (%)**	**Identity (%)**	**E-Value**	**GB Accession**	**Species$^{(*)}$**
C13	9755	2633	100	96.55	0	LN611622.1	AbMV
C56	254.375	2617	100	98.89	0	KJ939791.1	BGMV
C163	254.375	2617	100	98.89	0	KJ939791.1	
C236	2205	2645	96	99.09	0	NC_015397.1	CleGMV
C89	13.826	2630	96	91.21	0	NC_015397.1	
C29	21.408	2697	96	81.17	0	MZ019476.1	EuSLGMV
C40	354.538	2609	100	98.47	0	JN419000.1	EuYMV
C153	346.256	2658	100	96.73	0	JN419000.1	
C48	354.538	2609	100	98.47	0	JN419000.1	
C3087	354.538	2609	100	98.47	0	JN419000.1	
C12	170.811	2674	86	82.28	0	JN419005.1	MacYSV
C238	356.819	2659	100	86.60	0	FN436001	SiBV
C124	769.084	2606	100	84.78	0	NC_014446.1	SiGMFlV
C63	102.231	2676	100	95.22	0	KC706535.1	SiMMV
C81	45.586	2676	100	96.00	0	KC706535.1	
C58	203.671	2676	100	96.04	0	KC706535.1	
C45	50.531	2645	100	96.40	0	KX348157.1	
C88	63.201	2611	90	86.02	0	JX415194.1	
C34	315.060	2676	100	96.11	0	KC706535.1	
C85	221.589	2669	100	93.78	0	KC706535.1	
C323	1,314.760	2667	99	91.49	0	AJ557450.1	
C55	126.06	2691	98	92.42	0	KC706535.1	
C30	730.419	2676	100	93.20	0	KC706535.1	
C32	58.900	2676	100	97.53	0	KC706535.1	
C186	1,636.588	2681	100	95.41	0	AJ557450.1	
C154	701.809	2676	100	93.12	0	KC706535.1	
C66	720.708	2676	100	92.79	0	KC706535.1	
C43	912.187	2612	100	90.81	0	KU852503.1	

C86	1,491.559	2680	100	93.59	0	AJ557450.1	
C98	1,319.979	2685	100	90.49	0	AJ557450.1	
C46	902.73	2618	100	89.71	0	KU852503.1	
C490	1,166.118	2664	100	91.67	0	AJ557450.1	
C835	788.863	2676	100	92.38	0	KU852503.1	
C79	861.612	2616	92	87.34	0	MT103998.1	SiYBV
C75	599.178	2649	100	90.55	0	MT103991.1	
C80	42.682	2678	100	84.24	0	NC_020255.1	SiYMAV
C162	14.652	2666	100	86.66	0	NC_020255.1	
C150	602.718	2671	100	86.18	0	MF957207.1	SiYNV
C16	1,087.484	2671	100	87.00	0	MF957207.1	
C60	526.946	2720	100	91.70	0	JX871374.1	ToMMV
C8	693.675	2720	100	94.30	0	JX871374.1	
C52	693.975	2720	100	94.30	0	JX871374.1	
C62	564.450	2719	100	92.87	0	JX871374.1	
C53	659.941	2687	100	94.29	0	JX871374.1	
C131	801.031	2680	100	89.93	0	JX871374.1	
C136	3.262	2593	100	98.96	0	MW596594.1	ToSRV
C7	215.212	2646	97	95.12	0	MT214096.1	ToYNV
C2340	595.131	2656	98	82.00	0	MN518741.1	

*AbMV: Abutilon mosaic virus; BGMV: Bean golden mosaic virus; CleGMV: Cleome golden mosaic virus; EuSLGMV: Euphorbia severe leaf golden mosaic virus; EuYMV: Euphorbia yellow mosaic virus; MacYSV: Macroptilium yellow mosaic virus; SiBV: Sida Brasil virus; SiGMFlV: Sida golden mosaic Florida 1 virus; SiMMV: Sida micrantha mosaic virus; SiYBV: Sida yellow blotch virus; SiYMAV: Sida yellow mosaic Alagoas virus; SiYNV: Sida yellow net virus; ToMMV: Tomato mild mosaic virus; ToSRV: Tomato severe rugose virus and ToYNV: Tomato yellow net virus. Putative new species are highlighted and gray.

Table 2 – Number of contigs, reading coverage, assembled genome size, BLASTn coverage, sequence identity of the assembled virus, E-value, GenBank (GB) accession of begomoviruses discovery by High-Throughput Sequencing (HTS) in a pool of 91 foliar samples of weeds (from 11 botanic families), containing DNA–B segments of distinct *Begomovirus* species.

Contigs	Coverage reads	Size (nucleotides)	Coverage (%)	Identity (%)	E-Value	GB accession	Begomovirus species
C175	27.948	2622	100.00	87.84	0	LN611625.1	AbMV
C2	382.438	2593	100.00	99.04	0	MN822293.1	BGMV
C21	360.638	2593	100.00	99.08	0	MN822293.1	
C220	548.498	2572	88.00	99.08	0	MN822293.1	
C1242	172.565	2656	85.00	90.80	0	MN822293.1	
C193	20.794	2675	100.00	92.04	0	MT627024.1	BlYSV
C37	457.711	2578	100.00	97.79	0	JN419001.1	EuYMV
C228	296.810	2552	86.00	92.36	0	JN419001.1	
C125	51.781	2591	94.00	88.71	0	MT104028.1	PSLDV
C64	66.650	2613	100.00	88.94	0	FN436002.1	SiBV
C465	2450.00	2587	100.00	83.11	0	HM359016.1	SiGMFlV
C320	853.000	2572	100.00	95.54	0	NC_014447.1	SiGMFlV
C6059	343.700	2605	100.00	89.98	0	NC_014447.1	
C24	81.854	2614	100.00	81.70	0	NC_016579.1	SiMALV
C35	820.539	2614	100.00	81.01	0	NC_016579.1	
C23	75.576	2641	100.00	80.80	0	NC_016579.1	
C137	34.630	2675	76.00	77.47	0	NC_015044.1	SiMBoV1
C4	3.681.328	2628	100.00	91.08	0	FN557523.1	SiMMV
C42	89.801	2656	100.00	94.39	0	KC706533.1	
C9	224.708	2619	89.00	79.40	0	HM585434.1	
C57	89.801	2656	100.00	94.39	0	KC706533.1	
C111	87.426	2615	100.00	92.88	0	KY650723.1	
C65	71.542	2654	100.00	91.66	0	KY650723.1	
C107	51.758	2639	100.00	91.90	0	KY650716.1	
C41	68.895	2669	100.00	94.27	0	AJ557452.1	
C11	71.158	2660	100.00	94.07	0	KX691416.1	
C54	3.681.328	2628	100.00	91.08	0	FN557523.1	SiMMV
C68	99.267	2616	86.00	87.09	0	MT104007.1	
C2734	612.890	2651	86.00	93.84	0	KC706533.1	

C448	1.147.791	2694	90.00	90.99	0	KC706533.1	
C112	75.772	2657	100.00	93.60	0	KY650723.1	
C61	119.133	2590	98.00	93.34	0	KC706533.1	
C16	1.087.484	2671	100.00	87.00	0	MF957207.1	SiYNV
C10	150.889	2603	100.00	79.47	0	NC_055471.1	ToCLCV
C47	306.376	2621	100.00	76.39	0	NC_055471.1	
C28	364.168	2706	82.88	66.00	0	KC706609.1	TOMIMV
C39	1.202.657	2703	100.00	92.73	0	KC706609.1	
C18	1.384.317	2675	100.00	93.18	0	KC706609.1	
C1087	1.221.334	2700	100.00	92.46	0	KC706609.1	
C20	1.462.074	2675	100.00	93.29	0	KC706609.1	
C834	1.362.911	2600	92.00	94.27	0	KC706613.1	
C4516	1.873.873	2645	88.00	89.01	0	KC706609.1	
C50	1.350.091	2659	100.00	92.73	0	KC706609.1	
C3	1.354.813	2676	100.00	93.48	0	KC706609.1	
C19	1.341.191	2676	100.00	93.59	0	KC706609.1	
C134	4.734	2570	100.00	97.20	0	MG837739.1	ToSRV

*AbMV: Abutilon mosaic virus; BGMV: Bean golden mosaic virus BlYSV: Blainvillea yellow spot virus; EuYMV: Euphorbia yellow mosaic virus; PSLDV: Passionfruit severe leaf distortion virus; SiBV: Sida Brasil virus; SiGMFlV: Sida golden mosaic Florida 1 virus; SiMALV: Sida mosaic Alagoas virus; SiMBoV1: Sida mosaic Bolivia virus; SiMMV: Sida micrantha mosaic virus; SiYNV: Sida yellow net virus; ToCLCV: Tomato chlorotic leaf curl virus; ToMMV: Tomato mild mosaic virus and ToSRV: Tomato severe rugose virus. Putative new species are highlighted and gray.

Table 3 – Number of contigs, reading coverage, assembled genome size, BLASTn coverage, sequence identity of the assembled virus, E-value, GenBank accession of geminiviruses and subviral agents discovery by High-Throughput Sequencing (HTS) in a pool of weed foliar samples, containing six isolates (distinct contigs) of the genera *Clecrusatellite*. *Mulcrilevirus*, and *Topilevirus*.

Genera	Contigs	Coverage reads	Size	Coverage (%)	Identity (%)	E-Value	GenBank	Viral species and subviral agent^(*)^
Clecrusatellite	C151	8578	1337	100	96.12	0	KY559642.1	EYMA
Mulcrilevirus	C97	2981	2968	94	92.25	0	MN240483.1	MCAV
Topilevirus	C791	3388	2873	100	97.91	0	MT135209.1	ToALCV
	C393					0		
	C708					0		
	C1259					0		

* EYMA: Euphorbia yellow mosaic alphasatellite; MCAV: Mungbean curl associated virus and ToALCV: tomato apical leaf curl virus.

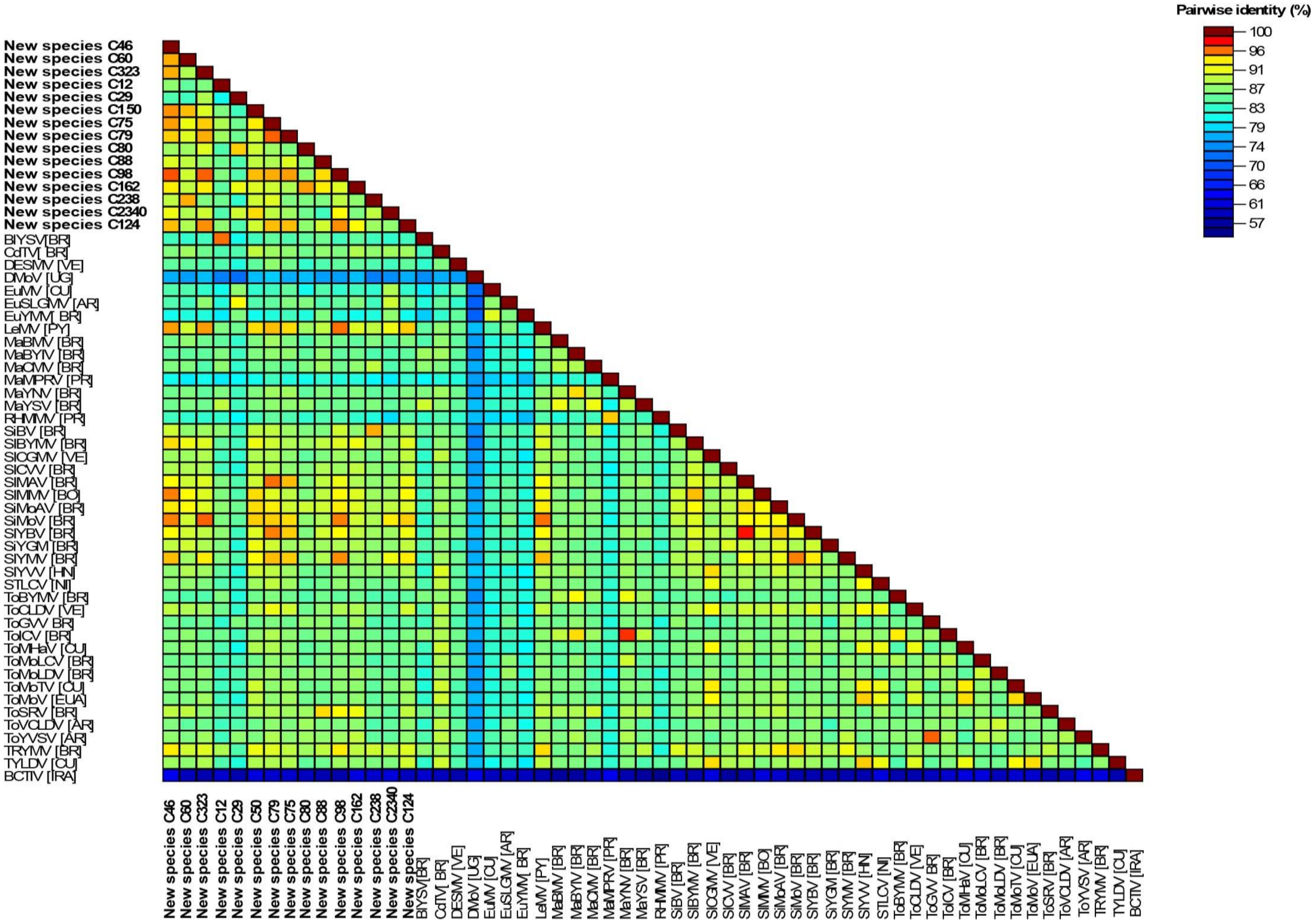

Pairwise identity (%)
100
96
91
87
83
79
74
70
66
61
57
New species C46
New species C60
New species C323
New species C12
New species C29
New species C150
New species C75
New species C79
New species C80
New species C88
New species C98
New species C162
New species C238
New species C2340
New species C124
BlYSV[BR]
CdTV[BR]
DESMV [VE]
DMoV [UG]
EuMV [CU]
EuSLGMV [AR]
EuYMV[BR]
LeMV [PY]
MaBMV [BR]
MaBYIV [BR]
MaCMV [BR]
MaMPRV [PR]
MaYNV [BR]
MaYSV [BR]
RHMMV [PR]
SiBV [BR]
SIBYMV [BR]
SICGMV [VE]
SICVV [BR]
SIMAV [BR]
SIMMV [BO]
SiMoAV [BR]
SiMoV [BR]
SIYBV [BR]
SiYGM [BR]
SIYMV [BR]
SIYVV [HN]
STLCV [NI]
ToBYMV [BR]
ToCLDV [VE]
ToGVV BR]
ToICV [BR]
ToMHaV [CU]
ToMoLCV [BR]
ToMoLDV [BR]
ToMoTV [CU]
ToMoV [EUA]
ToSRV [BR]
ToVCLDV [AR]
ToYVSV [AR]
TRYMV [BR]
TYLDV [CU]
BCTIV [IRA]
New species C46
New species C60
New species C323
New species C12
New species C29
New species C50
New species C79
New species C75
New species C80
New species C88
New species C98
New species C162
New species C238
New species C2340
New species C124
BlYSV[BR]
CdTV[BR]
DESMV [VE]
DMoV [UG]
EuMV [CU]
EuSLGMV [AR]
EuYMV[BR]
LeMV [PY]
MaBMV [BR]
MaBYIV [BR]
MaCMV [BR]
MaMPRV [PR]
MaYNV [BR]
MaYSV [BR]
RHMMV [PR]
SiBV [BR]
SIBYMV [BR]
SICGMV [VE]
SICVV [BR]
SIMAV [BR]
SIMMV [BO]
SiMoAV [BR]
SiMoV [BR]
SIYBV [BR]
SIYGM [BR]
SIYMV [BR]
SIYVV [HN]
STLCV [NI]
ToBYMV [BR]
ToCLDV [VE]
ToGVV BR]
ToICV [BR]
ToMHaV [CU]
ToMoLCV [BR]
ToMoLDV [BR]
ToMoTV [CU]
ToMoV [EUA]
ToSRV [BR]
ToVCLDV [AR]
ToYVSV [AR]
TRYMV [BR]
TYLDV [CU]
BCTIV [IRA]

Figure 1 – Sequence Demarcation Tool (SDT) pair-by-pair identity analysis performed using DNA-A sequence information from *Begomovirus* species obtained from the NCBI database. These species are identified by acronym. by the abbreviations of the countries where they were described: AG=Argentina; BO= Bolivia; BR= Brazil; CU= Cuba; USA= United States of America; HN= Honduras; IR= Iran; NI= Nigeria; PR= Puerto Rico; PY= Paraguay; UG=Uganda; VE= Venezuela. Viral species and GenBank accession numbers are as follows: Sida mosaic Alagoas virus – SIMAV (BR- JF694471); Sida yellow mosaic virus – SIYMV (BR- JX871380); New species C75; New species C79; Sida yellow mosaic virus – SIYMV(BR- BR-AY090558); New species C12; Sida mottle virus – SiMoV (BR- AY090555); New species C323; Leonurus mosaic virus – LeMV (KC683374.1); New species C98; New species C46; New species C60; New species C50; Sida bright yellow mosaic virus – SIBYMV (BR-KX348184); Sida micrantha mosaic virus – SIMMV (BO- HM585433); Sida Brazil virus – SiBV (BR- NC_038887.1); New species C238; Triumfetta yellow mosaic virus – TRYMV(BR- KX348184); Sida mottle Alagoas virus – SiMoAV (BR- JX871385); New species C162; New species C80; tomato severe rugose virus – ToSRV (BR- MW573997.1); New species C88; Sida chlorotic vein virus – SICVV (BR- Y11099); Sida yellow golden mosaic virus – SiYGM (BR- MF163258); Euphorbia mosaic virus – EuMV (CU- HQ896201); Euphorbia yellow mosaic virus – EuYM (BR- MF163258); Euphorbia severe leaf golden mosaic virus – EuSLGMV (AR-MZ019476); New species C162; New species C234; Sida ciliaris golden mosaic virus – SICGMV (VE-JX857691);); tomato yellow distortion leaf virus – TYLDV (CU-HE806438); tomato mottle Taino virus – ToMoTV (CU-AF012300); tomato mottle virus – ToMoV (USA-L14460); Sida yellow vein virus – SIYVV (HN-Y11099) ; tomato chlorotic leaf distortion virus – ToCLDV (VE-JN241632); tomato leaf curl Sinaloa virus – STLCV (NI-AJ608286); tomato mosaic Havana virus – ToMHaV (CU-Y14874); Chino del tomato Amazonas virus – CdTV (BR-MH243423); Macroptilium mosaic Puerto Rico virus – MaMPRV (PR-AY044133); Rhynchosia mild mosaic virus – RHMMV (PR-FJ944019); Desmodium mottle virus – DMoV (UG-KY294724); tomato golden vein virus – ToGVV(BR-JF803254); tomato yellow vein streak virus – ToYVSV (AR-MN508216); tomato mottle leaf curl virus – ToMoLCV (BR-KX896398); tomato mottle leaf distortion virus – ToMoLDV (BR-MW561191); tomato vein clearing leaf deformation virus – ToVCLDV (AR-MK4232208); Desmodium mosaic virus – DESMV (VE-OK044474); Blainvillea yellow spot virus – BlYSV (BR-EU710756); New species C12; Macroptilium yellow spot virus – MaYSV (BR-JN419013); Macroptilium common mosaic virus – MaCMV (BR-KX691396); Macroptilium bright mosaic virus – MaBMV (BR-KX691399); Macroptilium yellow net virus – MaYNV (BR-JN418998); tomato interveinal chlorosis virus – ToICV (BR-JF803252); Macroptilium bright yellow interveinal virus – MaBYIV (BR-MN146017); tomato bright yellow mosaic virus – ToBYMV (BR-KC791690), and Beet curly top Iran virus – BCTIV (IRA-JX082259).

Figure 2 – Middle root Bayesian phylogenetic tree with GTR+I model) for DNA–A components of the selected begomoviruses. These species are identified by acronym. by the abbreviations of the countries where they were described: AG=Argentina; BO= Bolivia; BR= Brazil; CU= Cuba;

USA= United States of America; HN= Honduras; IR= Iran; NI= Nigeria; PY= Paraguay; UG=Uganda; VE= Venezuela. Viral species and GenBank accession numbers are as follows: Sida mosaic Alagoas virus – SIMAV (BR- JF694471.); Sida yellow mosaic virus – SIYMV (BR- JX871380); New species C75; New species C79; Sida yellow mosaic virus – SIYMV(BR- BR-AY090558); New species C12; Sida mottle virus – SiMoV (BR - AY090555); New species C323; Leonurus mosaic virus – LeMV (KC683374.1); New species C98; New species C46; New species C60; New species C50; Sida bright yellow mosaic virus – SIBYMV (BR- KX348184); Sida micrantha mosaic virus – SIMMV (BO- HM585433); Sida Brazil virus – SiBV (BR - NC_038887.1); New species C238; Triumfetta yellow mosaic virus – TRYMV(BR - KX348184); Sida mottle Alagoas virus – SiMoAV (BR- JX871385); New species C162; New species C80; tomato severe rugose virus – ToSRV (BR- MW573997.1); New species C88; Sida chlorotic vein virus – SICVV (BR- Y11099); Sida yellow golden mosaic virus – SiYGM (BR- MF163258); Euphorbia mosaic virus – EuMV (CU- HQ896201); Euphorbia yellow mosaic virus – EuYM (BR- MF163258); Euphorbia severe leaf golden mosaic virus – EuSLGMV (AR-MZ019476); New species C162; New species C234; Sida ciliaris golden mosaic virus – SICGMV (VE-JX857691);); tomato yellow distortion leaf virus – TYLDV (CU-HE806438); tomato mottle Taino virus – ToMoTV (CU-AF012300); tomato mottle virus – ToMoV (USA-L14460); Sida yellow vein virus – SIYVV (HN-Y11099) ; tomato chlorotic leaf distortion virus – ToCLDV (VE-JN241632); tomato leaf curl Sinaloa virus – STLCV (NI-AJ608286); tomato mosaic Havana virus – ToMHaV (CU-Y14874); Chino del tomato Amazonas virus – CdTV (BR-MH243423); Macroptilium mosaic Puerto Rico virus – MaMPRV (PR-AY044133); Rhynchosia mild mosaic virus – RHMMV (PR-FJ944019); Desmodium mottle virus – DMoV (UG-KY294724); tomato golden vein virus – ToGVV(BR-JF803254); tomato yellow vein streak virus – ToYVSV (AR-MN508216); tomato mottle leaf curl virus – ToMoLCV (BR-KX896398); tomato mottle leaf distortion virus – ToMoLDV (BR-MW561191); tomato vein clearing leaf deformation virus – ToVCLDV (AR-MK4232208); Desmodium mosaic virus – DESMV (VE-OK044474); Blainvillea yellow spot virus – BlYSV (BR-EU710756); New species C12; Macroptilium yellow spot virus – MaYSV (BR-JN419013); Macroptilium common mosaic virus – MaCMV (BR-KX691396); Macroptilium bright mosaic virus – MaBMV (BR-KX691399); Macroptilium yellow net virus – MaYNV (BR-JN418998); tomato interveinal chlorosis virus – ToICV (BR-JF803252); Macroptilium bright yellow interveinal virus – MaBYIV (BR-MN146017); tomato bright yellow mosaic virus – ToBYMV (BR-KC791690) and Beet curly top Iran virus – BCTIV (IRA-JX082259).

REFERENCES

1. Fiallo-Olivé, E., Lett, J. M., Martin, D. P., Roumagnac, P., Varsani, A., Zerbini, F. M., ... & Consortium, I. R. (2021). ICTV virus taxonomy profile: *Geminiviridae* 2021. *The Journal of General Virology*, 102.
2. ICTV. *International Committee on Taxonomy of Viruses*. Available at: https://talk.ictvonline.org/taxonomy/. Accessed 06 April 2023.
3. Azzam, O., Frazer, J., Rosa, D., Beaver, J.S., Ahlquist, P., Maxwell D.P. (1994) Whitefly transmission and efficient ssDNA accumulation of Bean golden mosaic geminivirus require functional coat protein. *Virology*, 204,289-296.
4. Briddon R.W., Pinner M.S., Stanley J., Markham P.G. (1990) Geminivirus coat protein gene replacement alters insect specificity. *Virology*, 177,85-94.
5. Frischmuth, T., Stanley, J. (1998) Recombination between viral DNA and the transgenic coat protein gene of African cassava mosaic geminivirus. *Journal of General Virology*, 79,1265-1271.
6. Harrison, B.D., Swanson, M.M., Fargette, D. (2002) *Begomovirus* coat protein: Serology, variation and functions. *Physiological and Molecular Plant Pathology*, 60,257-271.
7. Fiallo-Olivé, E., Pan, L.L., Liu, SS., & Navas-Castillo, J. (2020) Transmission of Begomoviruses and Other Whitefly-Borne Viruses: Dependence on the Vector Species. *Phytopathology*, 110,10-17.
8. Padidam, M., Beachy, RN., Fauquet, CM. (1996) The role of AV2 ("precoat") and coat protein in viral replication and movement in tomato leaf curl geminivirus. *Virology*, 224,390-404.
9. Laufs, J., Jupin, I., David, C., Schumacher, S., Heyraud-Nitschke, F., Gronenborn, B. (1995) Geminivirus replication: genetic and biochemical characterization of Rep protein function, a review. *Biochimie*, 77,765-773.
10. Bisaro, D.M. (1996) Geminivirus DNA replication. Replication in Eukaryotic Cells, M. DePamphilis (Eds.) Cold Spring Harbor Laboratory Press (833-854 pp). Cold Spring Harbor, NY.
11. Sunter, G., & Bisaro, D.M. (1991). Transactivation in a geminivirus: AL2 gene product is needed for coat protein expression. *Virology*, 180,416-419.
12. Bisaro, D.M. (2006) Silencing suppression by geminivirus proteins. *Virology*, 344,158-168.
13. Sunter, G., Hartitz, M.D., Hormuzdi, S.G., Brough, C.L., & Bisaro, D.M. (1990). Genetic analysis of tomato golden mosaic virus: ORF AL2 is required for coat protein accumulation while ORF AL3 is necessary for efficient DNA replication. *Virology*, 179, 69-77.
14. Jupin, I., De Kouchkovsky, F., Jouanneau, F., Gronenborn, B. (1994) Movement of Tomato yellow leaf curl geminivirus (TYLCV): involvement of the protein encoded by ORF C4. *Virology*, 204,82-90.
15. Hanley-Bowdoin, L., Bejarano, E.R., Robertson, D., Mansoor, S. (2013) Geminiviruses: masters at redirecting and reprogramming plant processes. *Nature reviews. Microbiology*, 11,777-788.
16. Li, F., Xu, X., Huang, C., Gu, Z., Cao, L., Hu, T., Ding, M., Li, Z., Zhou, X. (2015) The AC5 protein encoded by Mungbean yellow mosaic India virus is a

pathogenicity determinant that suppresses RNA silencing-based antiviral defenses. *The New Phytologist*, 208,555-569.

17. Noueiry, A.O., Lucas, W.J., Gilbertson, R.L. (1994) Two proteins of a plant DNA virus coordinate nuclear and plasmodesmal transport. *Cell*, 76,925-932.
18. Hehnle, S., Wege, C., Jeske, H. (2004) Interaction of DNA with the movement proteins of geminiviruses revisited. *Journal of Virology*, 78,7698-706.
19. Li, H., Jiang, Z., Zhou, J., Liu, X., Zhang, Y., & Chu, D. (2023). Ecological factors associated with the distribution of *Bemisia tabaci* cryptic species and their facultative endosymbionts. *Insects*, 14, 252.
20. Dinsdale, A., Cook, L., Riginos, C., Buckley, Y.M, De Barro, P. (2010). Análise global refinada de *Bemisia tabaci* (Hemiptera: Sternorrhyncha: Aleyrodoidea: Aleyrodidae) mitocondrial citocromo oxidase 1 para identificar limites genéticos em nível de espécie. *Annals of the Entomological Society of America*, 103,196-208.
21. Fiallo-Olivé, E., & Navas-Castillo, J. (2023). Begomoviruses: what is the secret (s) of their success?. *Trends in Plant Science*, 1360-1385.
22. Giordano, L.B., Fonseca, M. E. N., Silva, J. B., Inoue-Nagata, A. K., & Boiteux, L. S. (2005). Efeito da infecção precoce por *Begomovirus* com genoma bipartido em características de frutos de tomate industrial. *Horticultura Brasileira*, *23*, 815-818.
23. Fontenele, R. S., Ribeiro, G. C., Lamas, N. S., Ribeiro, S. G., Costa, A. F., Boiteux, L. S., & Fonseca, M. E. N. (2018). First report of sida micrantha mosaic virus infecting Oxalis species in Brazil. *Plant disease*, *102*(9), 1862.
24. Duarte, MF., Fonseca, MEN., Boiteux, LS., Costa, H., Ribeiro, BM., Melo, FL., Pereira-Carvalho, RC. (2020) Identification of *Physalis angulata* (Solanaceae) as a natural alternative weed host of tomato severe rugose virus in Brazil. *Plant Disease*, 104,600-600.
25. Duarte, M. F., Fonseca, M. E., Costa, H., Fernandes, N. A., Reis, A., Boiteux, L. S., & Pereira-Carvalho, R. C. (2021a). Diversity of tomato-infecting begomoviruses and spatiotemporal dynamics of an endemic viral species of the Brazilian Atlantic rain forest biome. *Virus Genes*, *57*, 83-93.
26. Duarte, M. F., Pereira-Carvalho, R. D. C., Reis, L. N. A., Rojas, M. R., Gilbertson, R. L., Costa, H., Boiteux, L.S. & Fonseca, M. E. N. (2021b). Natural infection of tomatoes (*Solanum lycopersicum*) by Euphorbia yellow mosaic virus isolates across four Brazilian states. *Plant Disease*, *105*(2), 518.
27. Reis, LNA, Fonseca, M. E. N., Ribeiro, S. G., Naito, F. Y. B., Boiteux, L. S., & Pereira-Carvalho, R.C. (2020). Metagenomics of neotropical single-stranded DNA viruses in tomato cultivars with and without the *Ty*-1 gene. *Viruses*, *12*(8), 819.
28. Reis, LNA, Boiteux, L.S., Fonseca, & Pereira-Carvalho, R.C. (2021). Tomato yellow vein streak virus and tomato golden vein virus: A reappraisal of the classification status of two South American Begomovirus species based upon genome–wide pairwise identity of multiple isolates. *Virus Genes*, *57*, 127-131.
29. Pereira-Silva, J., Boiteux, L. S., Fonseca, M. E. N., Reis, L. N. A., Souza, A. S., Nery, F. M. B., Madeira, N. R. & Pereira-Carvalho, R. C. (2022). Novel natural hosts of tomato severe rugose virus (ToSRV) in the Fabaceae, Solanaceae, and Oxalidaceae families. *Journal of Plant Diseases and Protection*, 1-7.

30. Rodríguez-Negrete, E.A.; Morales-Aguilar, J.J.; Domínguez-Duran, G.; Torres-Devora, G.; Camacho-Beltrán, E.; Leyva-López, N.E.; Voloudakis, A.E.; Bejarano, ER; Méndez-Lozano, J. (2019) High-Throughput Sequencing reveals differential *Begomovirus* species diversity in non-cultivated plantas in northern pacific Mexico. *Viruses,* 11,594.
31. Rivarez, M.P.S., Pecman, A., Bačnik, K., Olivera M., Vučurović, A., Seljak G., Mehle N., Gutiérrez-Aguirre I., Ravnikar, M., Kutnjak D. (2023) In-depth study pf tomato and weed viromes reveals undiscovered plant virus diversity in na agroecosystem. *Microbiome*, 11,60.
32. Batista, J. G., Melo, F. L., Pereira-Carvalho, R. C., Alves-Freitas, D. M. T., & Ribeiro, S. G. (2019). First report of tomato apical leaf curl virus infecting tomato in Brazil. *Plant Disease,* 103, 1443.
33. Ma, Y., Navarro, B., Zhang, Z., Lu, M., Zhou, X., Chi, S., Di Serio, F., Li, S. (2015) Identification and molecular characterization of a novel monopartite geminivirus associated with Mulberry mosaic dwarf disease. *Journal of General Virology,* 96,2421-2434.
34. Boiteux, L.S., Fonseca, M.E.N., & Simon, P.W. (1999) Effects of plant tissue and DNA purification method on randomly amplified polymorphic DNA-based genetic fingerprinting analysis in carrot. *Journal of the American Society for Horticultural Science,* 124,32-38.
35. Inoue-Nagata, A.K., Albuquerque, L., Rocha, W., Nagata, T. (2004) A simple method for cloning the complete begomovirus genome using the bacteriophage φ29 DNA polymerase. *Journal of Virological Methods*, 116,209-211.
36. Rojas, M.R., Gilbertson, R., Russell, D., Maxwell, D. (1993) Use of degenerate primers in the polymerase chain reaction to detect whitefly-transmitted geminiviruses. Plant Disease, 77: 340-347 specificity. *Virology,* 177,85-94.
37. Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28,1647-1649.
38. Guindon, S., Lethiec, F., Duroux, P., & Gascuel, O. (2005). PHYML Online-a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Research*, 33,557-559.
39. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, 28, 2731-2739.
40. Hu T, Chitnis N, Monos D, Dinh A. (2021) Next-generation sequencing technologies: An overview. *Human Immunology,* 82,801-811.
41. Qiu, Y., Zhang, S., Yu, H., Xuan, Z., Yang, L., Zhan, B., Murilo Zerbini F., Cao, M. (2020) Identification and characterization of two novel geminiviruses associated with paper mulberry (*Broussonetia papyrifera*) leaf curl disease. *Plant Disease,* 104,3010-3018.
42. Lu, Q.Y., Wu, Z.J., Xia, Z., Xie, L.H. (2015) Complete genome sequence of a novel monopartite geminivirus identified in mulberry (*Morus alba* L.). *Archives of Virology,* 160,2135-2138.

43. Nogueira, A. M., Nascimento, M. B., Barbosa, T. M. C., Quadros, A. F. F., Gomes, J. P. A., Orílio, A. F., Barros, D. R., & Zerbini, F. M. (2021). The Association between New World Alphasatellites and Bipartite Begomoviruses: Effects on Infection and Vector Transmission. *Pathogens*, 10,1244.
44. Mar, T. B., Mendes, I. R., Lau, D., Fiallo-Olivé, E., Navas-Castillo, J., Alves, M. S., & Murilo Zerbini, F. (2017). Interaction between the New World begomovirus *Euphorbia yellow mosaic virus* and its associated alphasatellite: effects on infection and transmission by the whitefly *Bemisia tabaci*. *Journal of General Virology*, 98,1552-1562.
45. Nawaz-ul-Rehman, M. S., Nahid, N., Mansoor, S., Briddon, R. W., & Fauquet, C. M. (2010). Post-transcriptional gene silencing suppressor activity of two non-pathogenic alphasatellites associated with a begomovirus. *Virology*, 405, 300-308.
46. Paprotka, T., Metzler, V., Jeske, H. (2010) The first DNA 1-like alpha satellites in association with New World begomoviruses in natural infections. *Virology*, 404,148-157.
47. Jeske, H., Kober, S., Schäfer, B., Strohmeier, S. (2014) Circomics of Cuban geminiviruses reveals the first alpha-satellite DNA in the Caribbean. *Virus Genes*, 49,312-324.
48. Sung, Y.K., Coutts, R.H. (1995) Pseudorecombination and complementation between potato yellow mosaic geminivirus and tomato Golden mosaic geminivirus. *Journal of General Virology*, London, 76,2809-2815.
49. Andrade, E.C., Manhani, G.G., Alfenas, P.F., Calegario, R.F., Fontes, E.P.B., & Zerbini, F.M. (2006) Tomato yellow spot virus, a tomato-infecting begomovirus from Brazil with a closer relationship to viruses from Sida sp., forms pseudorecombinants with begomoviruses from tomato but not from Sida. *Journal of General Virology*, 87,3687-3696.
50. Brown, J.K., Zerbini, M., Navas-Castillo, J., Moriones, E., Ramos-Sobrinho, R,. Silva, J.C.F., Fiallo-Olivé, E., Briddon, RW., Hernández-Zepeda, C.,Idris, A., Malathi, VG., Martin, D.P., Rivera-Bustamante, R., Ueda, S., Varsani, A. (2015) Revision of *Begomovirus* taxonomy based on pairwise sequence comparisons. *Archives of Virology* 160,1593-1619.
51. Duffy, S., & Holmes, E. C. (2009). Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. *The Journal of General Virology*, 90, 1539-1547
52. Padidam, M., Sawyer, S., Fauquet, C.M. (1999) Possible emergence of new geminiviruses by frequent recombination. *Virology*, 265,218-225.
53. García-Arenal, F., e Zerbini, F. M. (2019). Life on the edge: geminiviruses at the interface between crops and wild plant hosts. *Annual Review of Phytopathology*, 6,411-433.
54. Roye, M.E., McLaughlin, W.A., Nakhla, M.K., & Maxwell, D.P. (1997). Genetic diversity among geminiviruses associated with the weed species *Sida* spp., *Macroptilium lathyroides*, and *Wissadula amplissima* from Jamaica. *Plant Disease*, 81,1251-1258.
55. Faria, J. C., & Maxwell, D. P. (1999). Variability in geminivirus isolates associated with *Phaseolus* spp. in Brazil. *Phytopathology*, 89,262-268.

56. Jovel, J., Reski, G., Rothenstein, D., Ringel, M., Frischmuth, T., & Jeske, H. (2004). Sida micrantha mosaic is associated with a complex infection of begomoviruses different from Abutilon mosaic virus. *Archives of Virology,* 149,829-841.
57. Castillo-Urquiza, G. P., Beserra, J. E., Jr, Bruckner, F. P., Lima, A. T., Varsani, A., Alfenas-Zerbini, P., & Murilo Zerbini, F. (2008). Six novel begomoviruses infecting tomato and associated weeds in Southeastern Brazil. *Archives of Virology,* 153,1985-1989.
58. Barbosa, J. C., Teixeira, L. D. D., & Rezende, J. A. M. (2010). First report on the susceptibility of sweet pepper crops to tomato chlorosis virus in Brazil. *Plant Disease,* 94, 374.
59. Da Silva, S. J., Castillo-Urquiza, G. P., Hora Júnior, B. T., Assunção, I. P., Lima, G. S., Pio-Ribeiro, G., Mizubuti, E. S., & Zerbini, F. M. (2011). High genetic variability and recombination in a begomovirus population infecting the ubiquitous weed *Cleome affinis* in northeastern Brazil. *Archives of Virology,* 156, 2205–2213.
60. Fernandes, F. R., Albuquerque, L. C., Oliveira, C. L., Cruz, A. R., Rocha, W. B., Pereira, T. G., Naito, F. Y., Dias, N.M., Nagata, T., Faria, J. C., Zerbini, F. M., Aragão, F. J., & Inoue-Nagata, A. K. (2011). Molecular and biological characterization of a new Brazilian begomovirus, euphorbia yellow mosaic virus (EuYMV), infecting *Euphorbia heterophylla* plants. *Archives of Virology,* 156,2063-2069.
61. Leke, W. N., Brown, J. K., Ligthart, M. E., Sattar, N., Njualem, D. K., & Kvarnheden, A. (2012). Ageratum conyzoides: a host to a unique begomovirus disease complex in Cameroon. *Virus Research,* 163,229-237.
62. Fonseca, M. E. N., Boiteux, L. S., Abreu, H., Nogueira, I., & Pereira-Carvalho, R. C. (2013). *Physalis angulata*: O new natural host of *tomato chlorosis virus* in Brazil. *Plant Disease,* 97,692
63. Fernandes-Acioli, N. A. N., Boiteux, L. S., Fonseca, M. E. N., Segnana, L. R. G., & Kitajima, E. W. (2014). Report of tomato yellow spot virus infecting *Leonurus sibiricus* in Paraguay and within tomato fields in Brazil. *Plant Disease,* 98,1445.
64. Barreto, S. S., Hallwass, M., Aquino, O. M., & Inoue-Nagata, A. K. (2013). A study of weeds as potential inoculum sources for a tomato-infecting begomovirus in central Brazil. *Phytopathology,* 103,436–444.
65. Naito, F. Y., Melo, F. L., Fonseca, M. E. N., Santos, C. A., Chanes, C. R., Ribeiro, B. M., Gilbertson R.L. Boiteux L.S. & Pereira-Carvalho, R.C. (2019). Nanopore sequencing of a novel bipartite New World begomovirus infecting cowpea. *Archives of virology, 164,* 1907-1910.
66. Batista, J.G., Nery, F.M.B., Melo, F.F.S., Malheiros, M.F., Rezende, D.V., Boiteux, L.S., Fonseca, M.E.N., de Miranda, B.E.C., Pereira-Carvalho, R.C. (2022) Complete genome sequence of a novel bipartite begomovirus infecting the legume weed *Macroptilium erythroloma*. *Archives Virology,* 167,1597-1602.
67. Faria, J.C. de Aragão, FJL., Souza, TLPO., Quintela, ED., Kitajima, EW., Ribeiro, S., da G. (2016) Golden mosaic of common beans in Brazil: Management with a transgenic approach. APS Features, Available at: http://ainfo.cnptia.embrapa.br/digital/ bitstream/item/150080/1/CNPAF-2016-APS.pdf. Accessed 06 April 2023.

68. Fernandes, F. R., Cruz, A. R., Faria, J. C., Zerbini, F. M., & Aragão, F. J. (2009). Three distinct begomoviruses associated with soybean in central Brazil. *Archives of Virology*,154,1567-1570.
69. Walz, D. M., de Freitas, A. T., de Novaes, T. G., de Oliveira Molina, R., Bianchini, A., & Bussulo, J. (2015). Identification of alternative hosts of the *Fabaceae* family to the bean golden mosaic. *Blucher Biochemistry Proceedings*, 1,184-187.
70. NCBI. 2023. National Center for Biotechnology Information. GenBank. Available at: https://www.ncbi.nlm.nih.gov/.
71. Fernandes-Acioli, N. A. N., Pereira-Carvalho, R. C., Fontenele, R. S., Lacorte, C., Ribeiro, S. G., Fonseca, M. E. N., & Boiteux, L. S. (2011). First report of *Sida micrantha mosaic virus* in *Phaseolus vulgaris* in Brazil. *Plant Disease*, 95,1196.
72. Hoffmann, L. V., Inoue-Nagata, A. K., Vaz, L. N. A., Barroso, P. A. V., & Faria, J. C. D. (2022). Identification of *Sida micrantha mosaic virus* as the causal agent of common mosaic in cotton in Goiás. *Summa Phytopathologica*, 47, 222-224.
73. Paprotka T, Metzler V, Jeske H (2010) The complete nucleotide sequence of a new bipartite begomovirus from Brazil infecting Abutilon. *Archives of Virology*, 155,813-816

Supplementary Table 1. Information about the families, host, code, symptoms, location and date of the 91 leaf samples of vegetables belonging to 11 botanical families of weeds associated with tomato cultivation used in this work.

Family	Host	Code	Symptoms	Site*	Year
Amaranthaceae	*Amaranthus* spp.	CE–037	Mosaic	Ibiapina–CE	2010
	Amaranthus spp.	CE–038	Mosaic	Ibiapina–CE	2010
	Amaranthus spp.	DF–072	Yellow spots	CNPH Gama–DF	2003
	Amaranthus spp.	DF–574	Mosaic	Rajadinha–DF	2014
Asteraceae	*Bidens pilosa*	BA–002	Golden mosaic	Arraial D'Ajuda-BA	2007
	Sonchus oleraceus	DF–011	Mosaic	CNPH Gama–DF	2003
	Sonchus oleraceus	DF–012	Mosaic	CNPH Gama–DF	2003
Brassicaceae	*Brassica juncea*	ES–293	Mosaic	Domingos Martins–ES	2018
Cleomaceae	*Cleome* spp.	BA–005	Yellowing and chlorosis	Praia do Espelho–BA	2007
Convolvulaceae	*Ipomoea purpurea*	ES–001	Leaves chlorotic	Venda Nova do Imigrante–ES	2011
Cucurbitaceae	*Cyclanthera pedata*	DF–514	Mosaic	Rajadinha–DF	2013
Euphorbiaceae	*Euphorbia heterophylla*	BA–009	Golden mosaic	Irecê–BA	2007
	Euphorbia heterophylla	DF–064	Golden mosaic	CNPH Gama–DF	2003
	Euphorbia heterophylla	GO–398	Golden mosaic	Flores de Goiás–GO	2008
Fabaceae	*Stylosanthes* spp.	DF–179	Mosaic	Núcleo Rural São José–DF	2005
	Glycine max	DF–520	Mosaic	Rajadinha–DF	2013
Lamiaceae	*Melissa officinalis*	DF–411	Golden mosaic	Gama–DF	2011
Malvaceae	*Gossypium hirsutum*	GO–056	Apical mosaic	Goiânia–GO	2003
	Hibiscus spp.	DF–805	Apical mosaic	CNPH Gama–RS	2022
	Hibiscus spp.	RS–069	Golden mosaic	Canguçu–RS	2013
	Hibiscus spp.	RAL–759	Mosaic	São Miguel da Barra–AL	2022
	Hibiscus spp.	RDF–774	Apical mosaic	CNPH Gama–DF	2022
	Hibiscus spp.	RDF–775	Mosaic	CNPH Gama–DF	2022
	Hibiscus spp.	RDF–776	Mosaic	CNPH Gama–DF	2022

Malvaceae	*Malva* spp.	AM–004	Golden mosaic	Iranduba–BA	2007
	Malva spp.	DF–040	Golden mosaic	Lago Azul–DF	2003
	Malva spp.	DF–152	Mosaic	Asa Norte–DF	2005
	Malva spp.	DF–336	Golden mosaic and leaf distortion	CNPH Gama–DF	2010
	Malva spp.	DF–399	Golden mosaic	CNPH Gama–DF	2011
	Malva spp.	DF–494	Golden mosaic	Rajadinha–DF	2012
	Malva spp.	DF–618	Mosaic	CNPH Gama–DF	2015
	Malva spp.	ES–073	Mosaic	São Roque–ES	2012
	Malva spp.	GO–074	Golden mosaic	Santo Antônio Descoberto–GO	2003
	Malva spp.	GO–175	Golden mosaic	Leopoldo Bulhões–GO	2003
	Malva spp.	GO–263	Severe golden mosaic	Goianápolis–GO	2003
	Malva spp.	GO–283	Golden mosaic	Pirinópolis–GO	2004
	Malva spp.	GO–429	Golden mosaic	Teresina de Goiás–GO	2009
	Malva spp.	MG–289	Golden mosaic	Pará de Minas–MG	2010
	Malva spp.	PE–005	Golden mosaic	Chã Grande–PE	2007
	Malva spp.	PE–010	Golden mosaic	Comancin de S. Félix–PE	2007
	Malva spp.	PE–044	Golden mosaic	Comancin de S. Félix–PE	2010
	Malva spp.	PE–052	Golden mosaic	Novo Cajueiro–PE	2010
	Malva spp.	PE–069	Mosaic	Senharó–PE	2010
	Malva spp.	PE–115	Mosaic	Pesqueira–PE	2012
	Malva spp.	PI–001	Mosaic	Picos–PI	2015
	Malva spp.	SP–108	Mosaic	Capão Bonito–SP	2011
	Malva spp.	TO–008	Golden mosaic	Gurupi–TO	2004
	Malva spp.	TO–009	Chlorotic and purple leaves	Gurupi–TO	2004
	Malva spp.	TO–023	Golden mosaic	Miracena–TO	2005
	Malva spp.	TO–038	Golden mosaic	Guarai–TO	2005
	Malva spp.	TO–048	Leaves chlorotic	Gurupi–TO	2007

Malvaceae	*Malva* spp.	TO–086	Golden mosaic	Araguaiana–TO	2008
	Malva spp.	TO–135	Golden mosaic	Formoso–TO	2008
	Malva spp.	TO–265	Mottled and yellowish	Chapada de Natividade–TO	2009
	Malva spp.	TO–272	Golden mosaic	Conceição–TO	2009
	Malva spp.	TO–302	Golden mosaic	Taquaruçu–TO	2009
	Malva spp.	TO–308	Mosaic	IFTO–TO	2010
	Malva spp.	GO–216	Golden mosaic	Morrinhos–GO	2003
	Malva spp.	SP–133	Chlorotic leaves	Tupã–SP	2012
	Malva spp.	SP–134	Golden mosaic	Tupã–SP	2012
	Malva spp.	SP–135	Golden mosaic	Tupã–SP	2012
	Abelmoschus esculentus	ES–177	Mosaic	Venda Nova do Imigrante–ES	2016
	Abelmoschus esculentus	ES–178	Mosaic	Venda Nova do Imigrante–ES	2016
	Abelmoschus esculentus	MA–001	Mosaic	Vargem Grande–MA	2016
	Abelmoschus esculentus	TO–041	Mosaic	Gurupi–TO	2007
	Abelmoschus esculentus	TO–054	Chlorosis	Taquaruçu–TO	2007
	Abelmoschus esculentus	TO–105	Chlorosis	São João (Palmas)–TO	2008
	Abelmoschus esculentus	TO–313	Mosaic	Palmas–TO	2010
	Abelmoschus esculentus	DF–532	Mosaic	Ponte Alta–DF	2013
	Sida sp.	AM–017	Golden mosaic	Silves–AM	2007
	Sida sp.	BA–023	Golden mosaic	Irecê–BA	2007
	Sida sp.	BA–147	Mosaic	Irecê–BA	2011
Malvaceae	*Sida* sp.	CB–002	Golden mosaic	Caimito	2010
	Sida sp.	DF–802	Mosaic	CNPH Gama–DF	2022
	Sida sp.	MS–002	Mosaic	Campo Grande–MS	2019
	Sida sp.	PE–116	Mosaic	Venturosa–PE	2012
	Sida sp.	PR–017	Golden mosaic	São Gerônimo da Serra–PR	2007
	Sida sp.	PR–062	Asymptomatic	São Gerônimo da Serra–PR	2007

Malvaceae	*Sida* sp.	PR–063	Asymptomatic	São Gerônimo da Serra–PR	2007
	Sida sp.	RR–006	Mosaic	Boa Vista–RR	2016
	Sida sp.	TO–001	Golden mosaic	Alvorada–TO	2004
	Sida sp.	TO–035	Golden mosaic	Colméia–TO	2005
	Sida sp.	TO–078	Golden mosaic	Arraias–TO	2008
	Sida sp.	AC–004	Mosaic	Porto–AC	2016
	Sida sp.	RDF–828	Mosaic	Manhuaçu	2021
	Sida sp.	RDF–829	Mosaic	Manhuaçu	2021
	Sida sp.	PR–077	Golden mosaic	Reserva–PR	2008
	Sida sp.	RAL–760	Golden mosaic	São Miguel da Barra–AL	2022
Solanaceae	*Solanum paniculatum*	BA–016	Apical yellowing	Irecê–BA	2007
	Solanum macrocarpon	DF–668	Apical mosaic	CNPH Gama–DF	2016
	Solanum muricatum	DF–498	Mosaic	CNPH Gama–DF	2013

*AM: Amazonas; BA: Bahia; DF: Distrito Federal; ES: Espírito Santo; GO: Goiás; MA: Maranhão; MG: Minas Gerais; MS: Mato Grosso do Sul; PE: Pernambuco; PI: Piauí; PR: Paraná; RR: Roraima.

CHAPTER 3

Complete genome characterization of a novel recombinant *Begomovirus* species infecting a new natural Fabaceae host (*Bolusafra bituminosa*)

To be submitted to *Archives of Virology* – Annotated Sequence

Complete genome characterization of a novel recombinant bipartite *Begomovirus* species infecting a new natural Fabaceae host (*Bolusafra bituminosa*).

Eduardo Soares da Silva Lima[1], Juliana Pereira da Silva[1], Luciane de Nazaré Almeida dos Reis[1], Felipe Fochat S Melo[1], Leonardo S. Boiteux[1,2], Maria Esther de Noronha Fonseca[2], Rita de Cássia Pereira–Carvalho[1]

[1]Departamento de Fitopatologia, Universidade de Brasília (UnB), Brasília–DF, Brazil.

[2]National Center for Vegetable Crops Research (CNPH), Embrapa Vegetable Crops (Hortaliças), Brasília–DF, Brazil.

Abstract

Members of the genus *Begomovirus* (*Geminiviridae*) can infect a wide array of dicotyledonous hosts. Here, a new begomovirus was isolated from the weed *Bolusafra bituminosa* (Fabaceae) in Northeast Brazil. The complete sequence of the DNA–A component was recovered using high-throughput sequencing and further characterized via Sanger dideoxy-sequencing. This novel recombinant species was phylogenetically related to New World begomoviruses and the name Bolusafra interveinal chlorosis virus (BoICV) was proposed. The identification of BoICV substantiates the astonishing diversity of begomoviruses in Latin America and reinforces the importance of Fabaceae weeds as sources of potential new pathogens to major Fabaceae crops.

The family *Geminiviridae* (Order: *Geplafuvirales*) is composed of viruses with single-stranded DNA (ssDNA) genomes. This taxonomic group comprises 14 genera *Becurtovirus*, *Begomovirus*, *Capulavirus*, *Citlodavirus*, *Curtovirus*, *Eragrovirus*, *Grablovirus*, *Maldovirus*, *Mastrevirus*, *Mulcrilevirus*, *Opunvirus*, *Topilevirus*, *Topocuvirus*, and *Turnucurtovirus*. Host range, vector species, genomic organization, and phylogenetic relationships are the main taxonomic criteria for genus demarcation in *Geminiviridae* [1, 2, 3, 4]. The *Begomovirus* genus is the largest group within the *Geminiviridae* family [1], displaying a cosmopolitan distribution encompassing all temperate, subtropical, and tropical regions [5]. The genomes of the begomoviruses have either a monopartite (with only the DNA–A component) or a bipartite organization (with both DNA–A and DNA–B components), being transmitted by a complex of cryptic whitefly (Hemiptera: Aleyrodidae) species including *Bemisia tabaci* Middle East-Asia Minor 1 (= biotype B) [6, 7]. The transmission of the begomoviruses is predominantly non-propagative, persistent-circulative, although recent reports have indicated a propagative replication of some monopartite viruses in the salivary glands of *B. tabaci* [5, 8]. Begomoviruses can infect a wide array of dicotyledonous plants [3], including many Fabaceae hosts [9]. We carried out a survey in the virus database of the NCBI (https://www.ncbi.nlm.nih.gov/) that indicated a total of 144 begomoviruses infecting weed and crop species of the Fabaceae family **(Supplementary Table 1)**. In the present work, we ratify the diversity of begomoviruses infecting Fabaceae members by describing a novel virus species in a novel host: the invasive African weed *Bolusafra bituminosa* (L.) Kuntze (tribe Phaseoleae, Fabaceae).

A combination of two methodologies was employed to survey the viral diversity in Fabaceae weed plants displaying begomovirus-like symptoms (mainly golden mosaic) occurring in the North-East region of Brazil. Potential viruses present in these field-collected samples were initially recovered via high-throughput sequencing (HTS) and then further characterized via Sanger dideoxy-sequencing with the purpose of obtaining potential complete genomes of begomoviruses in these hosts. Twenty-three foliar samples of Fabaceae species were subjected to total DNA purification using the modified Cetyl Trimethyl Ammonium Bromide (CTAB) protocol with organic solvents [10]. The purified DNA was stored at -20°C and subsequently employed as a template in Rolling-Cicle Amplification (RCA) assays [11]. The initial confirmation of the presence of begomovirus infection in these samples was done by employing PCR (Polymerase Chain Reaction) tests with the degenerate primers 'PAR1c496' and 'PAL1v1978' [12]. Primer walking was then used to recover the complete sequence of the DNA–A genome [13]. For the initial detection of the DNA–B component, the primer pair 'PBL1v2040' / 'PCRc1' [12] was used. For HTS, circular DNAs of viral origin were enriched by RCA assays. HTS was then performed on Illumina NovaSeq 6000 platform. The HTS-derived information was analyzed with a previous described three-step workflow [14] as follows: ***(i)*** elimination of low-quality reads; ***(ii)*** re–assembly of the sequences using the program CLC Genomics Workbench 23.0.2; and ***(iii)*** BLASTn validation of the contigs by comparing with the ssDNA virus database of the GenBank (https://www.ncbi.nlm.nih.gov/). The viral contigs were annotated and the trimmed reads were mapped back to the annotated genome using the tool 'Map to reference' available in the Geneious 11.0 program [15]. Sequence identities to the closest begomoviruses were determined with Species Demarcation Tool v.1.2 (SDT) [16]. Full-length genomes were aligned with MUSCLE multiple alignment program. Phylogenetic trees based on DNA–A genome alignments were generated by IQtree 2.2.0 with model Bayesian Information

Criterion: TIM3+F+R5 with 1,000 bootstrap replications in the Geneious program. Figures were elaborated with Adobe Illustrator CC and EvolView [17]. To detect potential recombination events, the software RDP 5 program [18] was used. Recombination events were considered consistent only if they were detected by at least four out of the seven methods implemented by the program.

As a result of the Illumina NovaSeq 6000 sequencing, 7477430 million reads were obtained. After assembly, in the CLC Genomics Workbench 23.0.2 program, 47646 contigs were obtained, and one of them correspondent to new species of virus described in this work. The DNA–A component of a putative new begomovirus species displayed 2697 nucleotides (nts) with typical organization of New World bipartite species. The highest identity of this virus (85.87%%) was with tomato bright yellow mottle virus (ToBYMV – BR-KC791690), sharing the same clade (**Figure 1A and Figure 2**). The DNA–A was detected using PCR assays with virus-specific primers in the sample codified as CE–076 obtained from the invasive weed *Bolusafra bituminosa*, which was collected in September of 2002 in Quixeré–CE. The name Bolusafra interveinal chlorosis virus (BoICV) is tentatively proposed for this new species. One segment of 755 nts of the DNA–B was obtained and displayed identity of 91.53% with tomato crinkle virus (ToCrV). The complete sequence of the DNA–B component was not recovered. The nonatucleotide sequence found in most of the geminiviruses (5'-TAATATTAC-3') was also detected in the DNA–A component of BoICV. In addition, our analyzes identified the iteron, GGGGT (Rep IRD = MPPPKRFKIS) [19] (**Figure 1B**). Five ORFs (open reading frames) were detected in the DNA–A component, one of them in the viral sense (AV1) with 756 nts coding the Coat Protein (CP). In the complementary sense, four ORFs were detected *viz.* replication associated protein gene with 1056 nts (AC1), trans-acting protein gene with 390 nts (AC2), replication enhancer gene with 399 nts (AC3) and symptom determinant gene (AC4) with 294 nts (**Figure 1A**).

Strong evidence of two recombination events was detected in the DNA–A component using the RDP5 program. The first event was verified by seven methods: RDP (3.412 x 10^{-14}), GENECONV (5.019 x 10^{-41}), BootSan (9.506 x 10^{-21}), MaxChi (3.054 x 10^{-25}), Chimaera (1.431 x 10^{-25}), SiScan (5.048 x 10^{-21}), and 3Seq (7.149 x 10^{-53}). In this first event, the major parental was tomato bright yellow mosaic virus (ToBYMV) and the minor parental was tomato severe rugose virus (ToSRV). Recombination encompassed the CP, REn and TrAP regions from the position 319 to 1445 (total size = 1127 nts). The second recombination event was detected by six methods: GENECONV (5.019 x 10^{-11}), BootSan ($9.506 x 10^{-15}$), MaxChi ($3.054 x 10^{-08}$), Chimaera ($1.431 x 10^{-09}$), SiScan ($5.048 x 10^{-06}$), and 3Seq ($7.149 x 10^{-16}$). In this second event, the major parent was ToBYMV and the minor parent was Sida bright yellow mosaic virus (SiBYMV). Recombination comprised the common region and the beginning of the Rep gene (total length =161 nts).

The identification and characterization of a new recombinant bipartite species substantiates the astonishing diversity of begomoviruses infecting crops and weeds in the Caribbean and Latin America regions, with Brazil being one of the major centers of diversity for the neotropical bipartite begomoviruses [20]. The detection of BoICV reinforces the importance of Fabaceae weeds as alternative as well as primary sources of begomoviruses [21, 22, 23]. In addition, begomovirus-infected Fabaceae weeds might also act as sources of genomic segments for the potential emergence of novel recombinant begomoviruses able to infect tomato as well as other Fabaceae hosts of economic relevance.

Author contributions: Conceptualization, E.S.S.L, J.P.S, L.N.A.R, L.S.B. M.E.N.F and R.C.P.C; Data curation, E.S.S.L, J.P.S, L.S.B and R.C.P.C; Formal analysis, E.S.S.L, L.N.A.R and R.C.P.C; Investigation, E.S.S.L, L.S.B, R.C.P.C Project administration, L.S.B and R.C.P.C; Supervision, M.E.N.F, L.S.B and R.C.P.C; Figure 1, L.N.A.R, and R.C.P.C, Figure 2, E.S.S.L, L.N.A.R, F.F.S.M, and R.C.P.C, Supplementary table 1 E.S.S.L, Writing-original draft, E.S.S.L, L.S.B, and R.C.P.C. All authors reviewed the manuscript.

Funding: This research was supported by grants, scholarships and fellowships from Embrapa, FAP-DF, and CNPq. This study was also financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil (CAPES) – Finance Code 001.

Declarations

Conflict of interest: The authors declare that they have no conflicts of interest.

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

Acknowledgments

This research was supported by grants, scholarships and post–doc fellowships from Embrapa, FAP–DF, CAPES, and CNPq.

References

1. ICTV (2023) International Committee on Taxonomy of Viruses. https://ictv.global/. Accessed 25 March 2023

2. Loconsole G, Önelge N, Potere O, Giampetruzzi A, Bozan O, Satar S, De Stradis A, Savino V, Yokomi R, Saponari M (2012) Identification and characterization of Citrus yellow vein clearing virus, a putative new member of the genus Mandarivirus. Phytopathology 102:1168–1175. https://doi.org/10.1094/PHYTO-06-12-0140-R
3. Rojas, MR, Macedo M, Maliano M, Soto-Aguilar M, Souza J, Briddon R, Kenyon L, Rivera Bustamante R, Zerbini F, Adkins S (2018) World management of geminiviruses. Annu Rev Phytopathol 56:637–677. https://doi.org/10.1146/annurev-phyto-080615-100327
4. Varsani A, Roumagnac P, Fuchs M, Navas-Castillo J, Moriones E, Idris A, Briddon RW, Rivera-Bustamante R, Zerbini FM, Martin DP (2017) Capulavirus and Grablovirus: Two new genera in the family *Geminiviridae*. Arch Virol 162:1819–1831. https://doi.org/10.1007/s00705-017-3268-6.
5. Fiallo-Olivé E, Let JM, Martin DP, Roumagnac P, Varsani A, Zerbini FM, & Navas-Castillo J (2021) ICTV Virus Taxonomy Profile: *Geminiviridae* 2021. J Gen Virol., 102(12), 001696. https://doi.org/10.1099/jgv.0.001696
6. Brown JK, Zerbini FM, Navas-Castillo J, Moriones E, Ramos-Sobrinho R, Silva JC, Fiallo-Olivé E, Briddon RW, Hernández-Zepeda C, Idris A (2015) Revision of Begomovirus taxonomy based on pairwise sequence comparisons. Arch Virol 160:1593–1619. https://doi.org/10.1007/s00705-015-2398-y
7. De Barro P, Liu S-S, Boykin L, Dinsdale A (2011) *Bemisia tabaci:* a statement of species status. Annu Rev Entomol 56:1–19. 10.1146/annurev-ento-112408-085504. https://doi.org/10.1146/annurev-ento-112408-085504
8. He Y-Z, Wang Y-M, Yin T-Y, Fiallo-Olivé E, Liu Y-Q, Hanley-Bowdoin L, Wang X-W (2020) A plant DNA virus replicates in the salivary glands of its insect vector via recruitment of host DNA synthesis machinery. Proc Natl Acad Sci USA 117:16928–16937. https://doi.org/10.1073/pnas.1820132117
9. Kitajima EW (2020) An annotated list of plant viruses and viroids described in Brazil (1926–2018). Biota Neotrop 20:e20190932 https://doi.org/10.1590/1676-0611-BN-2019-0932
10. Boiteux L, Fonseca M, Simon P (1999) Effects of plant tissue and DNA purification method on randomly amplified polymorphic DNA-based genetic fingerprinting analysis in carrot. J Am Soc Hort Sci 124:32–38. https://doi.org/10.21273/JASHS.124.1.32
11. Inoue-Nagata AK, Albuquerque L, Rocha W, Nagata T (2004) A simple method for cloning the complete begomovirus genome using the bacteriophage φ29 DNA polymerase. J Virol Methods 116:209–211. https://doi.org/10.1016/j.jviromet.2003.11.015
12. Rojas, MR, Gilbertson R, Russell D, Maxwell D (1993) Use of degenerate primers in the polymerase chain reaction to detect whitefly-transmitted geminiviruses. Plant Dis 77:340–347. https://agris.fao.org/agris-search/search.do?recordID=US9406522
13. Duarte MF, Fonseca ME, Costa H, Fernandes NA, Reis A, Boiteux LS, Pereira-Carvalho RC (2021) Diversity of tomato-infecting begomoviruses and spatiotemporal dynamics of an endemic viral species of the Brazilian Atlantic rain forest biome. Virus Genes 57:83-93. https://doi.org/10.1007/s11262-020-01812-x

14. Reis LNA, Fonseca MEN, Ribeiro SG, Naito FYB, Boiteux LS, Pereira-Carvalho RC (2020) Metagenomics of neotropical single-stranded DNA viruses in tomato cultivars with and without the *Ty*-1 gene. Viruses 12: 819.
15. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649. https://doi.org/10.1093/bioinformatics/bts199
16. Muhire BM, Varsani A, Martin DP (2014) SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. PloS One 9: e108277. https://doi.org/10.1371/journal.pone.0108277
17. He Zilong, Zhang H, Gao S, Lercher MJ, Chen W-H, Hu S (2016) Evolview v2: an online visualization and management tool for customized and annotated phylogenetic trees. Nucleic Acids Res 44:W236–W241. https://doi.org/10.1093/nar/gkw370
18. Martin DP, Varsani A, Roumagnac P, Botha G, Maslamoney S, Schwab T, Kelz Z, Kumar V, Murrell B (2021) RDP5: A computer program for analyzing recombination in, and removing signals of recombination from, nucleotide sequence datasets. Virus Evol 7: veaa087. https://doi.org/10.1093/ve/veaa087
19. Argüello-Astorga G, Ruiz-Medrano R (2001) An iteron-related domain is associated to Motif 1 in the replication proteins of geminiviruses: identification of potential interacting amino acid-base pairs by a comparative approach. Arch Virol 146:1465–1485. https://doi.org/10.1007/s007050170072
20. Chinnadurai C, Kollam M, Ramsubhag A, Jayaraman J (2022) Diversity of begomoviruses infecting crops and weeds in the Caribbean and Latin America. In *Geminivirus: Detection, Diagnosis and Management* (pp. 139–158). Academic Press.
21. Batista JG, Pereira-Carvalho RC, Malheiros MF, Rezende DV, Reis LNA, Fonseca MEN, Boiteux LS (2020) *Macroptilium erythroloma* (Fabaceae): A natural weed host of bean golden mosaic virus in Brazil. Plant Disease 104: 3270.
22. Batista JG, Nery FM, Melo FFS, Malheiros, MF, Rezende DV, Boiteux LS, Fonseca MEN, Miranda, BEC, Pereira-Carvalho RC (2022) Complete genome sequence of a novel bipartite begomovirus infecting the legume weed *Macroptilium erythroloma*. Arch Virol 167: 1597–1602.
23. Pereira-Silva J, Boiteux LS, Fonseca MEN, Reis LNA, Souza AS, Nery FMB, Madeira N, Pereira-Carvalho RC (2022) Novel natural hosts of tomato severe rugose virus (ToSRV) in the Fabaceae, Solanaceae, and Oxalidaceae families. J Plant Dis Prot 129: 425–431.

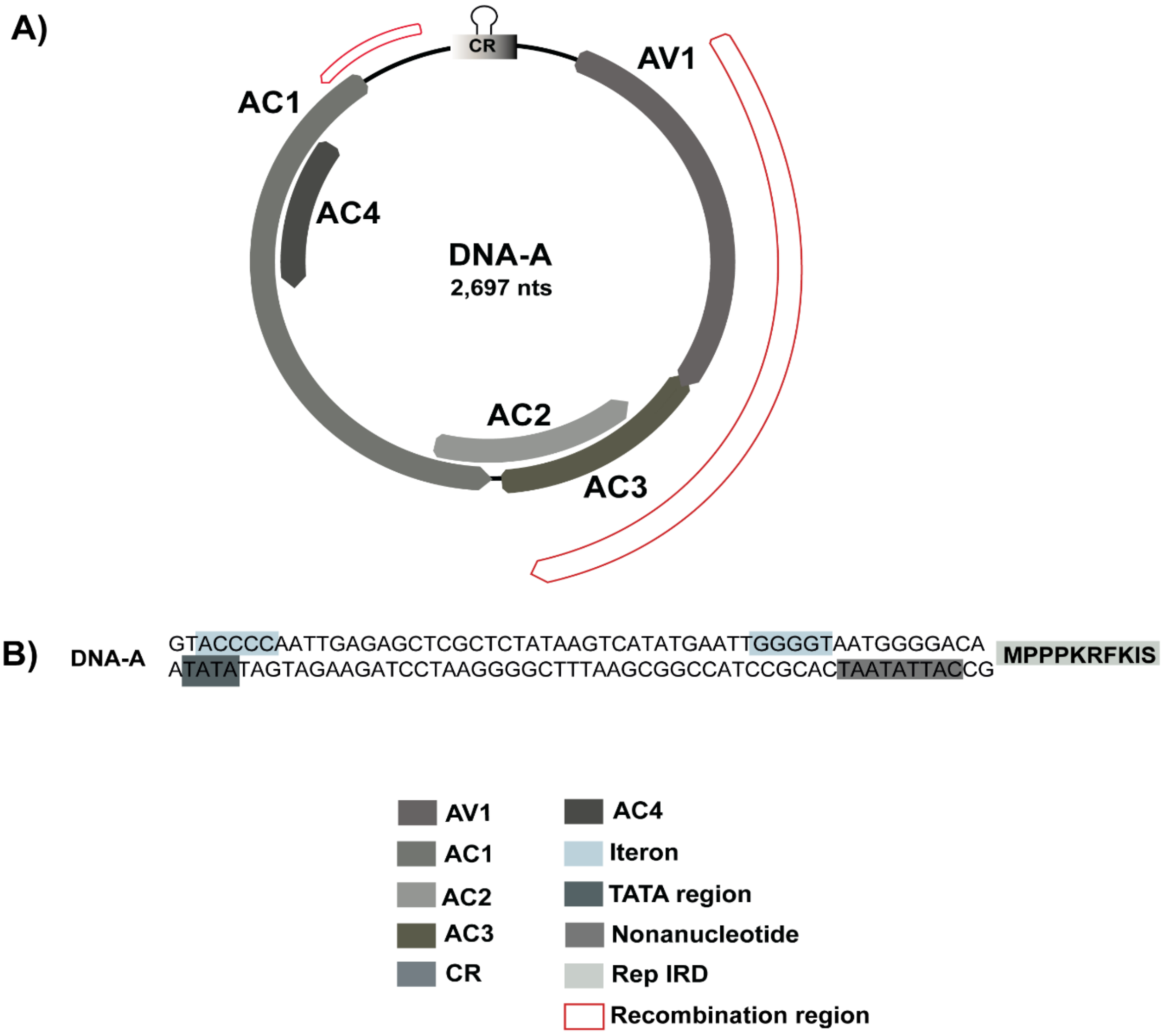
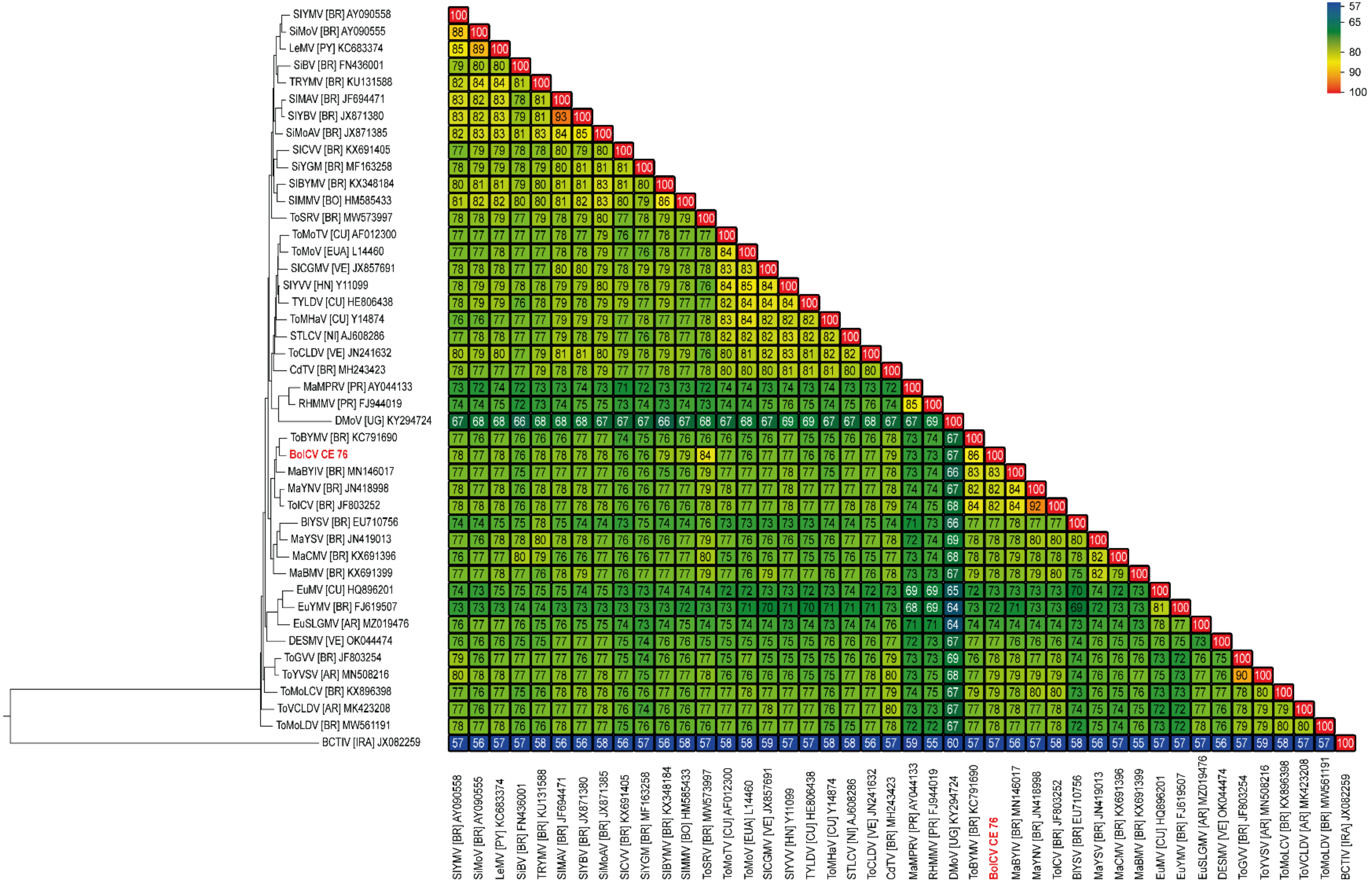


Figure 1. **(A)** The DNA–A component of the begomovirus isolate CE–076 (tentatively named as Bolusafra interveinal chlorosis virus – BoICV), displaying a typical genomic organization of New World species (i.e., devoid of the ORF AV2). Component DNA–A is composed of 2697 nucleotides presenting an intergenic region of 199 nucleotides at the stem-loop site. DNA–A component comprises five ORFs (open reading frames): AV1 (CP), AC1 (Rep), AC2 (TrAP), AC3 (REn), and AC4. **(B)** Details of the genomic region encompassing the Iteron, TATA region, nonanucleotide, and the Rep terminus = IRD (Rep Iteron–Related Domain).



	SIYMV [BR] AY090558	SiMoV [BR] AY090555	LeMV [PY] KC683374	SiBV [BR] FN436001	TRYMV [BR] KU131588	SIMAV [BR] JF694471	SIYBV [BR] JX871380	SiMoAV [BR] JX871385	SICVV [BR] KX691405	SIYGM [BR] MF163258	SIBYMV [BR] KX348184	SIMMV [BO] HM585433	ToSRV [BR] MW573997	ToMoTV [CU] AF012300	ToMoV [EUA] L14460	SICGMV [VE] JX857691	SIYVV [HN] Y11099	TYLDV [CU] HE806438	ToMHaV [CU] Y14874	STLCV [NI] AJ608286	ToCLDV [VE] JN241632	CdTV [BR] MH243423	MaMPRV [PR] AY044133	RHMMV [PR] FJ944019	DMoV [UG] KY294724	ToBYMV [BR] KC791690	**BolCV CE 76**	MaBYIV [BR] MN146017	MaYNV [BR] JN418998	ToICV [BR] JF803252	BIYSV [BR] EU710756	MaYSV [BR] JN419013	MaCMV [BR] KX691396	MaBMV [BR] KX691399	EuMV [CU] HQ896201	EuYMV [BR] FJ619507	EuSLGMV [AR] MZ019476	DESMV [VE] OK044474	ToGVV [BR] JF803254	ToYVSV [AR] MN508216	ToMoLCV [BR] KX896398	ToVCLDV [AR] MK423208	ToMoLDV [BR] MW561191	BCTIV [IRA] JX082259
SIYMV [BR] AY090558	100																																											
SiMoV [BR] AY090555	88	100																																										
LeMV [PY] KC683374	85	89	100																																									
SiBV [BR] FN436001	79	80	80	100																																								
TRYMV [BR] KU131588	82	84	84	81	100																																							
SIMAV [BR] JF694471	83	82	83	78	81	100																																						
SIYBV [BR] JX871380	83	82	83	79	81	93	100																																					
SiMoAV [BR] JX871385	82	83	83	81	83	84	85	100																																				
SICVV [BR] KX691405	77	79	79	78	78	80	79	80	100																																			
SiYGM [BR] MF163258	78	79	79	78	79	80	81	81	81	100																																		
SIBYMV [BR] KX348184	80	81	81	79	80	81	81	83	81	80	100																																	
SIMMV [BO] HM585433	81	82	82	80	80	81	82	83	80	79	86	100																																
ToSRV [BR] MW573997	78	78	79	77	79	78	79	80	77	78	79	79	100																															
ToMoTV [CU] AF012300	77	77	77	77	77	78	77	79	76	77	78	77	77	100																														
ToMoV [EUA] L14460	77	77	78	77	77	78	78	79	77	76	78	77	78	84	100																													
SICGMV [VE] JX857691	78	78	78	77	77	80	80	79	78	79	79	78	78	83	83	100																												
SIYVV [HN] Y11099	78	79	79	78	78	79	79	80	79	78	79	78	76	84	85	84	100																											
TYLDV [CU] HE806438	78	79	79	76	78	79	78	79	79	77	79	77	77	82	84	84	84	100																										
ToMHaV [CU] Y14874	76	76	77	77	77	79	79	79	77	78	77	77	78	83	84	82	82	82	100																									
STLCV [NI] AJ608286	77	78	78	77	77	79	78	79	77	76	78	77	78	82	82	82	83	82	82	100																								
ToCLDV [VE] JN241632	80	79	80	77	79	81	81	80	79	78	79	79	76	80	81	82	83	81	82	82	100																							
CdTV [BR] MH243423	78	77	77	77	79	78	78	79	77	77	78	77	78	80	80	80	81	81	81	80	80	100																						
MaMPRV [PR] AY044133	73	72	74	72	73	73	74	73	71	72	73	73	72	74	74	73	74	73	74	73	73	72	100																					
RHMMV [PR] FJ944019	74	74	75	72	73	74	75	75	73	74	73	74	73	74	74	75	76	75	74	75	76	74	85	100																				
DMoV [UG] KY294724	67	68	68	66	68	68	68	67	67	67	66	67	68	67	68	67	69	69	67	67	68	67	67	69	100																			
ToBYMV [BR] KC791690	77	76	77	76	76	76	77	77	74	75	76	76	76	76	75	76	77	75	75	76	76	78	73	74	67	100																		
BolCV CE 76	78	77	78	76	78	78	78	78	76	76	79	79	84	77	76	77	77	76	76	77	77	79	73	73	67	86	100																	
MaBYIV [BR] MN146017	77	77	77	75	77	77	77	77	76	76	75	76	79	77	77	77	78	77	75	77	77	78	73	74	66	83	83	100																
MaYNV [BR] JN418998	78	77	78	76	78	78	78	77	76	76	77	77	79	78	77	77	78	77	77	77	77	78	74	74	67	82	82	84	100															
ToICV [BR] JF803252	78	78	78	76	78	77	78	78	76	76	76	76	78	78	77	77	78	77	77	78	78	79	74	75	68	84	82	84	92	100														
BIYSV [BR] EU710756	74	74	75	75	78	75	74	74	73	75	74	74	76	73	73	73	73	73	74	74	75	74	71	73	66	77	77	78	77	77	100													
MaYSV [BR] JN419013	77	76	78	78	80	78	78	77	76	76	76	77	79	77	76	76	76	75	76	77	77	78	72	74	69	78	78	78	80	80	80	100												
MaCMV [BR] KX691396	76	77	77	80	79	76	76	77	76	76	77	77	80	75	76	76	77	75	75	76	75	77	73	74	68	78	78	79	78	78	78	82	100											
MaBMV [BR] KX691399	77	77	78	77	76	78	79	77	76	77	77	77	79	77	76	79	77	77	76	77	76	78	73	73	67	79	78	78	79	80	75	82	79	100										
EuMV [CU] HQ896201	74	73	75	75	75	75	74	75	73	73	74	74	74	72	72	73	73	72	73	74	72	73	69	69	65	72	73	73	73	73	70	74	72	73	100									
EuYMV [BR] FJ619507	73	73	73	74	74	73	73	73	73	73	73	72	73	73	71	70	71	70	71	71	71	73	68	69	64	73	72	71	73	73	69	72	73	73	81	100								
EuSLGMV [AR] MZ019476	76	77	77	76	75	76	75	75	74	73	74	75	74	75	74	74	73	74	73	74	74	76	71	71	64	74	74	74	74	74	73	74	74	73	78	77	100							
DESMV [VE] OK044474	76	76	76	75	75	77	77	77	75	74	76	76	75	75	75	75	75	75	76	76	76	77	72	73	67	77	77	76	76	76	74	76	75	75	76	75	73	100						
ToGVV [BR] JF803254	79	76	77	77	77	77	76	77	75	74	76	76	77	75	77	75	75	75	75	76	76	79	73	73	69	76	78	78	77	78	74	76	76	76	73	72	76	75	100					
ToYVSV [AR] MN508216	80	78	78	77	77	78	77	78	77	75	77	77	78	76	76	76	76	75	75	77	78	80	73	75	68	77	79	79	79	79	74	76	76	77	75	73	77	76	90	100				
ToMoLCV [BR] KX896398	77	76	77	75	76	78	78	78	76	76	77	76	77	77	78	77	77	77	76	77	77	79	74	75	67	79	79	78	80	80	73	76	75	76	73	73	77	77	78	80	100			
ToVCLDV [AR] MK423208	77	76	77	76	77	77	77	78	76	74	76	76	77	77	77	78	77	77	76	77	77	80	73	73	67	77	78	78	77	78	74	75	74	75	72	72	77	76	78	79	79	100		
ToMoLDV [BR] MW561191	78	77	78	76	76	78	77	77	75	76	76	76	76	77	76	77	77	77	75	75	77	78	72	72	67	77	78	77	77	78	72	75	74	76	73	72	78	76	79	79	80	80	100	
BCTIV [IRA] JX082259	57	56	57	57	58	56	56	58	56	57	56	58	57	58	58	59	57	57	58	58	56	57	59	55	60	57	57	56	57	58	58	56	57	55	57	57	57	56	57	59	58	57	57	100

Figure 2. Sequence Demarcation Tool (SDT) pair-by-pair identity analysis performed using DNA–A sequence information from *Begomovirus* species obtained from the NCBI database. Middle root Bayesian phylogenetic tree with GTR+I model) for DNA–A components Bolusafra interveinal chlorosis virus (BoICV). Paired identity analysis in SDT was performed using only the DNA–A component information. The panel shows the color coding for the paired identity levels between the DNA–A components of the selected begomoviruses. These species are identified by acronym, by the abbreviations of the countries where they were described: AG=Argentina; BO= Bolivia; BR= Brazil; CU= Cuba; USA= United States of America; HN= Honduras; IR= Iran; NI= Nigeria; PR= Puerto Rico; PY= Paraguay; UG=Uganda; VE= Venezuela. Viral species and GenBank accession numbers are as follows: Sida yellow mosaic virus – SIYMV (BR-AY090558); Sida mottle virus – SiMoV (BR-AY090555); Leonurus mosaic virus – LeMV (PY-KC683374); Sida Brazil virus – SiBV (BR-FN436001); Triumfetta yellow mosaic virus – TRYMV (BR-KU131588); Sida mosaic Alagoas virus – SIMAV (BR-JF694471); Sida yellow blot virus – SIYBV (BR-JX871380); Sida mottle Alagoas virus – SiMoAV (BR-JX871385); Sida chlorotic vein virus – SICVV (BR-KX691405); Sida yellow golden mosaic virus – SiYGM (BR-MF163258); Sida bright yellow mosaic virus – SIBYMV (BR-KX348184); Sida micrantha mosaic virus – SIMV (BO-HM585433); tomato severe rugose virus – ToSRV (BR-MW573997); tomato mottle Taino virus – ToMoTV (CU-AF012300); tomato mottle virus – ToMoV (USA-L14460); Sida ciliaris golden mosaic virus – SICGMV (VE-JX857691); Sida yellow vein virus – SIYVV (HN-Y11099); tomato yellow distortion leaf virus – TYLDV (CU-HE806438); tomato mosaic Havana virus – ToMHaV (CU-Y14874); tomato leaf curl Sinaloa virus – STLCV (NI-AJ608286); tomato chlorotic leaf distortion virus – ToCLDV (VE-JN241632); Chino del tomato Amazonas virus – CdTV (BR-MH243423); Macroptilium mosaic Puerto Rico virus – MaMPRV (PR-AY044133); Rhynchosia mild mosaic virus – RHMMV (PR-FJ944019); Desmodium mottle virus – DMoV (UG-KY294724); tomato bright yellow mosaic virus – ToBYMV (BR-KC791690); Macroptilium bright yellow interveinal virus – MaBYIV (BR-MN146017); Macroptilium yellow net virus – MaYNV (BR-JN418998); tomato interveinal chlorosis virus – ToICV (BR-JF803252); Blainvillea yellow spot virus – BlYSV (BR-EU710756); Macroptilium yellow spot virus – MaYSV (BR-JN419013); Macroptilium common mosaic virus – MaCMV (BR-KX691396); Macroptilium bright mosaic virus – MaBMV (BR-KX691399); Euphorbia mosaic virus – EuMV (CU-HQ896201); Euphorbia yellow mosaic virus – EuYM (BR-FJ619507); Euphorbia severe leaf golden mosaic virus – EuSLGMV (AR-MZ019476); Desmodium mosaic virus – DESMV (VE-OK044474); tomato golden vein virus – ToGVV(BR-JF803254); tomato yellow vein streak virus – ToYVSV (AR-MN508216); tomato mottle leaf curl virus – ToMoLCV (BR-KX896398); tomato vein clearing leaf deformation virus – ToVCLDV (AR-MK4232208); tomato mottle leaf distortion virus – ToMoLDV (BR-MW561191) and Beet curly top Iran virus – BCTIV (IRA-JX082259).

Supplementary Table 1. *Begomoviruses* reported in hosts belonging to the *Fabaceae* family.

Acesso	Host	Vírus species	Size	Local
MN158326	*Phaseolus vulgaris*	*Bean latent virus*	2566	Mexico
KX857726	*Phaseolus vulgaris*	*Bean leaf crumple virus*	2538	Colombia
MH469731	*Vigna unguiculata*	*Cowpea bright yellow mosaic virus*	2632	Brazil
MN414067	*Phaseolus vulgaris*	*Bean bushy stunt virus*	2572	Argentina
KX096982	*Phaseolus vulgaris*	*Common bean severe mosaic virus*	2577	Cuba
KY294727	*Desmodium* sp.	*Desmodium mottle virus*	2713	Uganda
EF016486	*Glycine max*	*Soybean blistering mosaic virus*	2605	Argentina
KC430936	*Sigmoidotropis elegans*	*Vigna yellow mosaic virus*	2602	Mexico
HM236368	*Rhynchosia minima*	*Rhynchosia golden mosaic Havana virus*	2609	Cuba
DQ406672	*Rhynchosia minima*	*Rhynchosia golden mosaic Sinaloa virus*	2578	Mexico
HM236371	*Rhynchosia minima*	*Rhynchosia rugose golden mosaic virus*	2555	Cuba
AM999981	*Rhynchosia minima*	*Rhynchosia yellow mosaic virus*	2740	Pakistan
KU569583	*Mucuna pruriens var.*	*Velvet bean golden mosaic virus*	2767	Taiwan
KX011473	*Phaseolus vulgaris*	*Common bean mottle virus*	2631	Cuba
EU339937	*Glycine max* sp.	*Rhynchosia golden mosaic virus*	2570	Mexico
KY196219	*Macroptilium* sp.	*Macroptilium golden yellow mosaic virus*	2631	Dominican Republic
KY001635	*Pisum sativum*	*Pea leaf distortion virus*	2738	Nepal
KX691396	*Macroptilium lathyroides*	*Macroptilium common mosaic virus*	2632	Brazil
KX691400	*Macroptilium lathyroides*	*Macroptilium bright mosaic virus*	2636	Brazil
KU852742	*Senna occidentalis*	*Senna leaf curl virus*	2742	India
KJ481205	*Lablab purpureus*	*Dolichos yellow mosaic virus*	2733	India
JQ283246	*Phaseolus vulgaris*	*Bean chlorotic mosaic virus*	2566	Venezuela
JN848772	*Phaseolus vulgaris*	*Bean white chlorosis mosaic virus*	2597	Venezuela
JN848771	*Phaseolus vulgaris*	*Bean chlorosis virus*	2632	Venezuela
JX122966	*Glycine max*	*Soybean chlorotic spot virus*	2585	Brazil
JQ866297	*Phaseolus vulgaris*	*French bean leaf curl virus*	2741	India
JN419002	*Centrosema brasilianum*	*Centrosema yellow spot virus*	2675	Brazil
JN419013	*Macroptilium lathyroides*	*Macroptilium yellow spot virus*	2660	Brazil
FJ944023	*Phaseolus vulgaris*	*Bean yellow mosaic Mexico virus*	2641	Mexico

FJ944019	*Rhynchosia minima*	*Rhynchosia mild mosaic virus*	2611	USA
FJ944020	*Rhynchosia minima*	*Rhynchosia mild mosaic virus*	2551	USA
HM777508	*Rhynchosia* sp.	*Rhynchosia yellow mosaic India virus*	2727	India
GQ472984	*Glycine max*	*Soybean mild mottle virus*	2768	Nigeria
GQ472985	*Glycine max*	*Soybean chlorotic blotch virus*	2708	Nigeria
FN543425	*Mucuna* sp.	*Velvet bean severe mosaic virus*	2733	India
FJ455449	*Crotalaria juncea*	*Sunn hemp leaf distortion virus*	2774	India
FJ792608	*Rhynchosia minima*	*Rhynchosia golden mosaic Yucatan virus*	2542	Mexico
EF585290	*Macroptilium lathyroides*	*Macroptilium yellow mosaic virus*	2630	Jamaica
DQ641691	*Pueraria montana*	*Kudzu mosaic virus*	2672	Viet Nam
DQ641695	*Mimosa* sp.	*Mimosa yellow leaf curl virus*	2757	Viet Nam
DQ875870	*Desmodium glabrum*	*Desmodium leaf distortion virus*	2569	Mexico
AJ627904	*Macrotyloma uniflorum*	*Horsegram yellow mosaic virus*	2728	India
MZ358127	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	958	Indonesia
MZ358133	*Vigna unguiculata*	*Tomato yellow leaf curl Kanchanaburi*	374	Indonesia
MZ088147	*Vigna mungo*	*Mungbean yellow mosaic virus*	190	India
MZ041264	*Medicago polymorpha*	*Tomato leaf curl Palampur virus*	640	India
MW791224	*Vigna radiata*	*Mungbean yellow mosaic India virus*	2661	India
MW791226	*Fabaceae*	*Mungbean yellow mosaic India virus*	2670	India
MW791233	*Vigna radiata*	*Horsegram yellow mosaic virus*	2669	India
MW791235	*Phaseolus vulgaris*	*Horsegram yellow mosaic virus*	2671	India
MW791236	*Glycine max*	*Horsegram yellow mosaic virus*	2668	India
MW816852	*Crotalaria juncea*	*Papaya leaf curl virus*	2758	India
ON461886	*Lablab purpureus*	*Dolichos yellow mosaic virus*	419	India
ON622515	*Vigna radiata*	*Mungbean yellow mosaic India virus*	553	India
MW814709	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	2746	India
MW814710	*Fabaceae*	*Mungbean yellow mosaic India virus*	2746	India
MW814711	*Vigna radiata*	*Mungbean yellow mosaic India virus*	2746	India
MW814713	*Vigna aconitifolia*	*Mungbean yellow mosaic virus*	2729	India
MW731690	*Lablab purpureus*	*Dolichos yellow mosaic virus*	2760	Bangladesh
MZ648028	*Trigonella foenum-*	*Tomato leaf curl Kerala virus*	2772	India

MZ648029	*Trigonella foenum-*	*Tomato leaf curl New Delhi virus*	2695	India
MW792462	*Glycine max*	*Mungbean yellow mosaic virus*	313	India
MW602393	*Pachyrhizus erosus*	*Tomato severe rugose virus*	1207	Brazil
OM643233	*Glycine max*	*Mungbean yellow mosaic India virus*	525	India
MW560615	*Pachyrhizus*	*Tomato severe rugose virus*	2593	Brazil
ON181435	*Pueraria montana var.*	*Kudzu mosaic virus*	2731	China
OL310477	*Phaseolus vulgaris*	*Papaya leaf curl China virus*	2735	China
OM256474	*Glycine max*	*Mungbean yellow mosaic India virus*	421	India
OK058527	*Phaseolus vulgaris*	*Bean golden yellow mosaic virus*	2647	Colombia
MW805421	*Pueraria montana var.*	*Kudzu mosaic virus*	2730	China
MW082628	*Phaseolus vulgaris*	*Bean golden yellow mosaic virus*	2644	Cuba
MW917145	*Vigna mungo*	*Mungbean yellow mosaic India virus*	2747	India
MW917146	*Vigna mungo*	*Mungbean yellow mosaic India virus*	2668	India
OK044469	*Vigna unguiculata*	*Cabbage leaf curl virus*	2535	Venezuela
OK044470	*Macroptilium bracteatum*	*Macroptilium mottle virus*	2632	Venezuela
OK044472	*Macroptilium bracteatum*	*Bean leaf crumple virus*	2600	Venezuela
OK044474	*Desmodium scorpiurus*	*Desmodium mosaic virus*	2629	Venezuela
OK044476	*Desmodium scorpiurus*	*Desmodium yellow spot virus*	2634	Venezuela
OK044478	*Rhynchosia minima*	*Cabbage leaf curl virus*	2589	Venezuela
OK044482	*Rhynchosia minima*	*Rhynchosia mottle virus*	2614	Venezuela
OK044484	*Vigna mungo*	*Cabbage leaf curl virus*	2575	Venezuela
OK044486	*Rhynchosia minima*	*Cabbage leaf curl virus*	2575	Venezuela
OK044494	*Vigna unguiculata*	*Cabbage leaf curl virus*	2575	Venezuela
MT920325	*Vigna aconitifolia*	*Mungbean yellow mosaic virus*	2737	Pakistan
MT920326	*Vigna aconitifolia*	*Cotton leaf curl Kokhran virus*	2748	Pakistan
MT944951	*Vigna radiata*	*Mungbean yellow mosaic India virus*	878	India
MT944961	*Fabaceae*	*Mungbean yellow mosaic India virus*	839	India
MT944964	*Glycine max*	*Mungbean yellow mosaic India virus*	730	India
MT944965	*Vigna radiata*	*Mungbean yellow mosaic India virus*	824	India
MW373471	*Glycine max*	*Mungbean yellow mosaic India virus*	1285	India
MW600934	*Glycine max*	*Mungbean yellow mosaic India virus*	2746	India

MZ235786	*Vigna mungo*	*Mungbean yellow mosaic India virus*	420	India
MZ235791	*Desmodium* sp.	*Mungbean yellow mosaic India virus*	420	India
MZ235792	*Vigna mungo*	*Mungbean yellow mosaic India virus*	2746	India
MZ356197	*Vigna mungo*	*Mungbean yellow mosaic virus*	2671	India
MZ485479	*Vigna unguiculata*	*Begomovirus mozlegume*	2740	Mozambique
MZ485482	*Glycine max*	*Begomovirus mozlegume*	2739	Mozambique
MZ485484	*Fabaceae*	*Begomovirus mozlegume*	2740	Mozambique
MZ485486	*Vigna unguiculata*	*Begomovirus mozlegume*	2717	Mozambique
MZ485489	*Glycine max*	*Begomovirus mozlegume*	2669	Mozambique
MZ485496	*Fabaceae*	*Begomovirus mozlegume*	2670	Mozambique
MT863331	*Vigna radiata*	*Mungbean yellow mosaic virus*	805	India
MT840348	*Vigna radiata*	*Mungbean yellow mosaic virus*	765	India
OK431079	*Vigna radiata*	*Mungbean yellow mosaic India virus*	2746	India
MT783246	*Vigna radiata var. radiata*	*Mungbean yellow mosaic India virus*	764	India
MZ501564	*Glycine max*	*Mungbean yellow mosaic virus*	525	India
MT671430	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	2746	India
MZ821026	*Dolichos* sp.	*Dolichos yellow mosaic virus*	2761	India
MZ202353	*Vigna radiata*	*Mungbean yellow mosaic virus*	774	India
MZ130504	*Vigna mungo*	*Mungbean yellow mosaic virus*	2655	India
MZ130505	*Vigna mungo*	*Mungbean yellow mosaic virus*	2655	India
MZ130506	*Vigna mungo*	*Mungbean yellow mosaic virus*	2655	India
MZ475993	*Vigna mungo*	*Mungbean yellow mosaic India virus*	829	India
MZ475998	*Desmodium* sp.	*Mungbean yellow mosaic India virus*	797	India
MT373116	*Abrus precatorius*	*Bidens pilosa leaf crumple virus*	2735	China
MT626860	*Phaseolus vulgaris*	*Bean golden mosaic virus*	2590	Brazil
MT626879	*Macroptilium lathyroides*	*Bean golden mosaic virus*	2586	Brazil
MT626892	*Phaseolus vulgaris*	*Bean golden mosaic virus*	2589	Brazil
MT626949	*Phaseolus lunatus*	*Bean golden mosaic virus*	2552	Brazil
MT627026	*Phaseolus vulgaris*	*Macroptilium yellow spot virus*	2611	Brazil
MT627030	*Macroptilium lathyroides*	*Macroptilium yellow spot virus*	2611	Brazil
MT627035	*Phaseolus lunatus*	*Macroptilium yellow spot virus*	2611	Brazil

MW436690	*Vigna mungo*	*Mungbean yellow mosaic virus*	931	India
MW436691	*Vigna mungo*	*Mungbean yellow mosaic virus*	880	India
MW436693	*Vigna mungo*	*Mungbean yellow mosaic virus*	880	India
MT232629	*Glycine max*	*Mungbean yellow mosaic India virus*	2746	India
MT232630	*Glycine max*	*Mungbean yellow mosaic India virus*	2661	India
MT276258	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	1204	India
MT276259	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	952	India
MT276262	*Vigna unguiculata*	*Tomato leaf curl New Delhi virus*	1097	India
MT276267	*Glycine max*	*Mungbean yellow mosaic India virus*	1296	India
MN895083	*Vigna aconitifolia*	*Cotton leaf curl Kokhran virus*	2748	Pakistan
MN910265	*Glycine max*	*Pedilanthus leaf curl virus*	2764	Pakistan
MN922309	*Vigna aconitifolia*	*Mungbean yellow mosaic virus*	2737	Pakistan
MN885463	*Glycine max*	*Mungbean yellow mosaic India virus*	2746	Pakistan
MN885477	*Glycine max*	*Mungbean yellow mosaic virus*	2737	Pakistan
MN885478	*Glycine max*	*Mungbean yellow mosaic virus*	2738	Pakistan
MN885479	*Glycine max*	*Mungbean yellow mosaic virus*	2738	Pakistan
MN885481	*Glycine max*	*Croton yellow vein virus*	2750	Pakistan
MN885482	*Glycine max*	*Pedilanthus leaf curl virus*	2758	Pakistan
MN885486	*Glycine max*	*Tomato leaf curl Kerala virus*	2765	Pakistan
MN885488	*Glycine max*	*Cotton leaf curl Kokhran virus*	2748	Pakistan
MN566097	*Senna occidentalis*	*Pedilanthus leaf curl virus*	2759	Pakistan
MN146017	*Macroptilium erythroloma*	*Macroptilium bright yellow interveinal*	2611	Brazil
MT300190	*Vigna mungo*	*Mungbean yellow mosaic India virus*	2767	India
MT318839	*Vigna mungo*	*Mungbean yellow mosaic virus*	830	India
MT350281	*Vigna mungo*	*Mungbean yellow mosaic India virus*	2765	India
MT027035	*Vigna mungo*	*Mungbean yellow mosaic India virus*	2746	India
MT027036	*Vigna mungo*	*Horsegram yellow mosaic virus*	2735	India
MT027038	*Vigna mungo*	*Mungbean yellow mosaic virus*	2729	India
MT027039	*Vigna mungo*	*Mungbean yellow mosaic virus*	2677	India
MK158209	*Albizia lebbeck*	*Pedilanthus leaf curl virus*	2757	Pakistan
MN318234	*Glycine max*	*Bean golden mosaic virus*	524	Brazil

MN803327	*Glycine max*	*Mungbean yellow mosaic India virus*	549	India
MN803328	*Glycine max*	*Mungbean yellow mosaic India virus*	549	India
MN734371	*Anadenanthera colubrina*	*Bean golden mosaic virus*	2617	Brazil
MN822293	*Macroptilium erythroloma*	*Bean golden mosaic virus*	2594	Brazil
MT161669	*Rhynchosia minima*	*Rhynchosia golden mosaic Sinaloa virus*	2584	Mexico
MN910266	*Glycine max*	*Cotton leaf curl Kokhran virus*	2748	Pakistan
MN166204	*Glycine max*	*Mungbean yellow mosaic India virus*	509	India
MT345791	*Macrotyloma uniflorum*	*Mungbean yellow mosaic virus*	951	India
KX216478	*Desmodium* sp.	*Passionfruit leaf distortion virus*	1117	Colombia
KX216480	*Rhynchosia minima*	*Rhynchosia golden mosaic Colombia*	2584	Colombia
KX216486	*Desmodium* sp.	*Rhynchosia golden mosaic Colombia*	1101	Colombia
KX246934	*Caesalpinia*	*Rhynchosia golden mosaic Colombia*	1101	Colombia
MT270285	*Vigna mungo*	*Mungbean yellow mosaic India virus*	918	India
MN414068	*Phaseolus vulgaris*	*Bean bushy stunt virus*	2522	Argentina
MG564658	*Phaseolus vulgaris*	*Pepper leaf curl Bangladesh virus*	2755	Pakistan
MT071114	*Vigna mungo*	*Mungbean yellow mosaic virus*	633	India
MT119671	*Fabaceae*	*Mungbean yellow mosaic virus*	687	India
MN166199	*M. atropurpureum*	*Bean latent virus*	898	Mexico
MN249689	*Fabaceae*	*Malvastrum yellow vein virus*	2750	China
MN218669	*Fabaceae*	*Tobacco leaf curl Yunnan virus*	2747	China
MT089918	*Vigna unguiculata*	*Tomato yellow leaf curl Shuangbai virus*	2749	China
MN020535	*Vigna radiata*	*Mungbean yellow mosaic India virus*	2748	India
MN020536	*Vigna radiata*	*Mungbean yellow mosaic India virus*	2672	India
MN158326	*Phaseolus vulgaris*	*Bean latent virus*	2566	Mexico
MN026272	*Vigna radiata*	*Mungbean yellow mosaic India virus*	951	India
MN602419	*Vigna radiata*	*Mungbean yellow mosaic virus*	2728	India
MN602421	*Vigna radiata*	*Horsegram yellow mosaic virus*	2671	India
MN602423	*Macrotyloma uniflorum*	*Horsegram yellow mosaic virus*	2735	India
MN602424	*Macrotyloma uniflorum*	*Horsegram yellow mosaic virus*	2670	India
MN698277	*Vigna mungo*	*Mungbean yellow mosaic India virus*	2746	India
MN698278	*Vigna mungo*	*Mungbean yellow mosaic virus*	2656	India

MN698281	*Vigna stipulacea*	*Mungbean yellow mosaic India virus*	2656	India
MN698282	*Vigna stipulacea*	*Horsegram yellow mosaic virus*	2735	India
MN698283	*Vigna stipulacea*	*Mungbean yellow mosaic virus*	2677	India
MN698285	*Vigna unguiculata*	*Horsegram yellow mosaic virus*	2735	India
MN698287	*Vigna aconitifolia*	*Horsegram yellow mosaic virus*	2735	India
MN698290	*Vigna aconitifolia*	*Mungbean yellow mosaic India virus*	2656	India
MN698296	*Vigna aconitifolia*	*Mungbean yellow mosaic virus*	2677	India
MN814423	*Vigna stipulacea*	*Mungbean yellow mosaic virus*	2729	India
MN814424	*Vigna stipulacea*	*Horsegram yellow mosaic virus*	2670	India
MK098184	*Glycine max*	*Mungbean yellow mosaic virus*	2736	Pakistan
LC487406	*Phaseolus vulgaris*	*Ageratum yellow vein virus*	2752	Japan
MK784271	*Fabaceae*	*Bitter gourd yellow vein virus*	1038	Bangladesh
MK784272	*Fabaceae*	*Bhendi yellow vein mosaic virus*	1043	Bangladesh
MN579642	*Glycine max*	*Mungbean yellow mosaic India virus*	1286	India
MH550115	*Trigonella foenum-*	*Pedilanthus leaf curl virus*	2784	Pakistan
MG551984	*Phaseolus vulgaris*	*Malvastrum yellow vein Lahore virus*	2739	Pakistan
MK317961	*Vigna radiata*	*Mungbean yellow mosaic virus*	2726	India
MK317962	*Vigna radiata*	*Mungbean yellow mosaic virus*	2679	India
MN062190	*Glycine max*	*Mungbean yellow mosaic India virus*	506	India
MK637450	*Macrotyloma uniflorum*	*Pedilanthus leaf curl virus*	545	India
MK174383	*C. mucunoides*	*Potato yellow mosaic Panama virus*	686	Panama
MK649930	*Clitoria ternatea*	*Mungbean yellow mosaic India virus*	506	India
LC481009	*Glycine max*	*Mungbean yellow mosaic India virus*	970	Indonesia
MK089808	*Phaseolus vulgaris*	*Sida golden mosaic virus*	2635	USA
MK618662	*Rhynchosia minima*	*Rhynchosia golden mosaic Sinaloa virus*	2578	Mexico
MK634355	*Rhynchosia minima*	*Rhynchosia golden mosaic virus*	2605	Mexico
MK087122	*Phaseolus vulgaris*	*French bean leaf curl Madikeri virus*	2753	India
MK241785	*Phaseolus vulgaris*	*Bean golden yellow mosaic virus*	2602	Mexico
MH885653	*Vigna radiata*	*Mungbean yellow mosaic India virus*	875	India
MK391940	*Glycine max*	*Horsegram yellow mosaic virus*	977	India
MK391941	*Vigna radiata*	*Horsegram yellow mosaic virus*	969	India

MK391947	*Macrotyloma uniflorum*	*Horsegram yellow mosaic virus*	1042	India
MK391949	*Tephrosia purpurea*	*Horsegram yellow mosaic virus*	1043	India
MK391950	*Senna tora*	*Horsegram yellow mosaic virus*	1027	India
MK409376	*Vigna mungo*	*Mungbean yellow mosaic virus*	1112	India
MH324445	*Glycine max*	*Mungbean yellow mosaic India virus*	2742	India
MH324446	*Glycine max*	*Mungbean yellow mosaic India virus*	2673	India
MH255791	*Vigna radiata*	*Mungbean yellow mosaic India virus*	2747	India
MH255792	*Vigna radiata*	*Mungbean yellow mosaic India virus*	2672	India
MH925107	*Phaseolus lunatus*	*Bean golden mosaic virus*	2552	Brazil
MH795972	*Lablab purpureus*	*Dolichos yellow mosaic virus*	2760	India
MH481901	*Phaseolus vulgaris*	*Pepper leafroll virus*	2580	Ecuador
MG677964	*Phaseolus vulgaris*	*Tomato yellow leaf curl virus*	543	China
MG659315	*Fabaceae*	*Bean golden yellow mosaic virus*	2644	Mexico
MF774563	*Vigna mungo*	*Cotton leaf curl Kokhran virus*	2750	Pakistan
MF774564	*Vigna radiata*	*Cotton leaf curl Kokhran virus*	2748	Pakistan
MG334552	*Phaseolus vulgaris*	*Bean golden mosaic virus*	2617	Brazil
MG544284	*Phaseolus vulgaris*	*Tomato leaf curl Joydebpur virus*	2770	India
MG970362	*Phaseolus vulgaris*	*Tomato yellow leaf curl virus*	2766	Oman
KX710156	*Cyamopsis tetragonoloba*	*Okra enation leaf curl virus*	2740	Pakistan
KX710160	*Sesbania bispinosa*	*Pedilanthus leaf curl virus*	2758	Pakistan
KX710161	*Sesbania bispinosa*	*Papaya leaf curl virus*	1396	Pakistan
KX710162	*Sesbania bispinosa*	*Tomato leaf curl New Delhi virus*	2359	Pakistan
MF693401	*Cajanus cajan*	*Mungbean yellow mosaic India virus*	2739	India
MF693402	*Cajanus cajan*	*Mungbean yellow mosaic India virus*	2672	India
MH359391	*Phaseolus vulgaris*	*Cabbage leaf curl virus*	2530	Ecuador
MH359395	*Rhynchosia minima*	*Cabbage leaf curl virus*	2506	Ecuador
MH359396	*Glycine max*	*Cabbage leaf curl virus*	2575	Ecuador
MF429941	*Melilotus*	*Tomato yellow leaf curl virus*	777	Iraq
MH181823	*Glycine max*	*Mungbean yellow mosaic India virus*	2749	India
MH181824	*Glycine max*	*Mungbean yellow mosaic India virus*	2671	India
MH165181	*Phaseolus vulgaris*	*Tobacco curly shoot virus*	2743	China

MH205950	*Phaseolus vulgaris*	*Tomato yellow leaf curl virus*	2781	China
MG963202	*Vigna unguiculata*	*Cowpea bright yellow mosaic virus*	527	Brazil
MF163536	*Vigna aconitifolia*	*Mungbean yellow mosaic India virus*	972	India
MF163537	*Vigna mungo*	*Mungbean yellow mosaic India virus*	948	India
MF163538	*Vigna radiata*	*Mungbean yellow mosaic India virus*	288	India
MF035965	*Cajanus cajan*	*Cajanus cajan yellow mosaic virus*	858	India
KY886918	*Phaseolus vulgaris*	*Bean golden yellow mosaic virus*	1411	Cuba
KX671564	*Glycine max*	*Tomato leaf curl Kerala virus*	2756	Pakistan
KX671565	*Glycine max*	*Mungbean yellow mosaic India virus*	2751	Pakistan
KY996454	*Vigna unguiculata*	*Tomato yellow leaf curl virus*	2754	Trinidad and
KX711621	*Glycine max*	*Mungbean yellow mosaic India virus*	2751	Pakistan
KX711622	*Glycine max*	*Pedilanthus leaf curl virus*	2756	Pakistan
KY769274	*Cyamopsis tetragonoloba*	*Tomato yellow leaf curl virus*	511	Pakistan
KU958512	*Vicia faba*	*Tomato yellow leaf curl virus*	537	Tunisia
KY509042	*Mucuna pruriens*	*Velvet bean severe mosaic virus*	544	India
MF374736	*Glycine max*	*Mungbean yellow mosaic India virus*	506	India
KX655575	*Glycine max*	*Mungbean yellow mosaic India virus*	913	India
KX655576	*Vigna radiata*	*Mungbean yellow mosaic India virus*	893	India
KX655577	*Lablab purpureus*	*Mungbean yellow mosaic India virus*	730	India
KX655578	*Cajanus cajan*	*Mungbean yellow mosaic India virus*	760	India
KX655579	*Fabaceae*	*Mungbean yellow mosaic India virus*	838	India
KY555798	*Fabaceae*	*Sida Brazil virus*	2664	Argentina
KY555800	*Fabaceae*	*Tomato mottle wrinkle virus*	2565	Argentina
KY555801	*Fabaceae*	*Tomato yellow vein streak virus*	2567	Argentina
KX452226	*Phaseolus vulgaris*	*Mungbean yellow mosaic India virus*	2746	Oman
LN849710	*Lens culinaris*	*Tomato leaf curl New Delhi virus*	2738	Pakistan
LC271790	*Vigna radiata*	*Mungbean yellow mosaic India virus*	2747	India
LC271792	*Glycine max*	*Mungbean yellow mosaic India virus*	2746	India
KX857725	*Phaseolus vulgaris*	*Bean leaf crumple virus*	2598	Colombia
KY824799	*Fabaceae*	*Mungbean yellow mosaic virus*	894	India
LN845918	*Cyamopsis tetragonoloba*	*Pedilanthus leaf curl virus*	1085	Pakistan

KX246859	*Cajanus cajan*	*Tomato leaf curl Karnataka virus*	2760	India
KU852503	*Glycine max*	*Sida micrantha mosaic virus*	2677	Brazil
KY640302	*Glycine max*	*Mungbean yellow mosaic India virus*	477	India
KX710157	*Cyamopsis tetragonoloba*	*Tomato yellow leaf curl virus*	2759	Pakistan
KY556679	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	2748	India
KY556680	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	2654	India
KX827599	*Glycine max*	*Tomato leaf curl New Delhi virus*	2739	Pakistan
LN845913	*Cyamopsis tetragonoloba*	*Papaya leaf curl virus*	2747	Pakistan
KX828624	*Glycine max*	*Tomato severe rugose virus*	2593	Brazil
KY303697	*Vigna radiata*	*Mungbean yellow mosaic virus*	2654	Bangladesh
KX185518	*Phaseolus vulgaris*	*Bean golden yellow mosaic virus*	2645	Cuba
LT009395	*Cyamopsis tetragonoloba*	*Papaya leaf curl virus*	2752	Pakistan
LT009399	*Cyamopsis tetragonoloba*	*Papaya leaf curl virus*	2751	Pakistan
KU514411	*Cyamopsis tetragonoloba*	*Papaya leaf curl virus*	2748	Pakistan
KX640991	*Phaseolus lunatus*	*Sida yellow blotch virus*	2660	Brazil
KU165788	*Glycine max*	*Euphorbia mosaic virus*	2609	Cuba
LN713272	*Sesbania*	*Pedilanthus leaf curl virus*	2758	Pakistan
KX011472	*Phaseolus vulgaris*	*Tobacco leaf curl Cuba virus*	2567	Cuba
KX011473	*Phaseolus vulgaris*	*Common bean mottle virus*	2631	Cuba
KX458238	*Phaseolus vulgaris*	*Tomato severe rugose virus*	2593	Brazil
KX691398	*Macroptilium lathyroides*	*Macroptilium yellow spot virus*	2658	Brazil
KT253639	*Cyamopsis tetragonoloba*	*Papaya leaf curl virus*	2756	India
KU500804	*Vigna unguiculata*	*Papaya leaf crumple virus*	2736	India
KU562964	*Glycine max*	*Tobacco yellow crinkle virus*	2601	Cuba
KX343909	*Cajanus scarabaeoides*	*Cowpea golden mosaic virus*	774	India
KT444617	*Centrosema pubescens*	*Soybean chlorotic blotch virus*	2671	Cameroon
KU145406	*Phaseolus vulgaris*	*Bean golden yellow mosaic virus*	2608	Cuba
KU950430	*Vigna radiata*	*Mungbean yellow mosaic India virus*	2741	India
KU950431	*Vigna radiata*	*Mungbean yellow mosaic India virus*	2656	India
KT779558	*Desmodium glabrum*	*Macroptilium yellow spot virus*	2627	Brazil
KU562963	*Phaseolus vulgaris*	*Tobacco leaf curl Cuba virus*	2610	Cuba

KP752088	*Phaseolus vulgaris*	*Horsegram yellow mosaic virus*	2735	India
KP779630	*Phaseolus vulgaris*	*Mungbean yellow mosaic India virus*	2745	India
LC114252	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	1422	Indonesia
KT944082	*Senna tora*	*Tomato leaf curl Patna virus*	2774	India
JQ327840	*Phaseolus lunatus*	*Pea leaf distortion virus*	2736	Nepal
JQ327845	*Phaseolus lunatus*	*Mungbean yellow mosaic India virus*	2729	Nepal
KT381193	*Glycine max*	*Rhynchosia golden mosaic Yucatan virus*	2581	Cuba
KT282129	*Lablab purpureus*	*Dolichos yellow mosaic virus*	1127	India
KT224368	*Glycine max*	*Chino del tomate virus*	708	Mexico
KT192632	*Glycine max*	*Rhynchosia golden mosaic Yucatan virus*	2581	Cuba
KT022086	*Glycine max*	*Rhynchosia golden mosaic virus*	696	Mexico
KT022087	*Glycine max*	*Pepper huasteco yellow vein virus*	668	Mexico
LC093232	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	1422	Indonesia
KR053204	*Macrotyloma uniflorum*	*Horsegram yellow mosaic virus*	1083	
KR053205	*Vigna mungo*	*Mungbean yellow mosaic India virus*	1095	
KR021041	*Vigna unguiculata subsp.*	*Mungbean yellow mosaic India virus*	1484	Indonesia
KP641345	*Rhynchosia*	*Tobacco yellow crinkle virus*	2591	Jamaica
KP641350	*Rhynchosia*	*Rhynchosia golden mosaic Yucatan virus*	2492	Jamaica
KP752090	*Phaseolus vulgaris*	*Rhynchosia yellow mosaic virus*	2741	India
KP828155	*Glycine max*	*Mungbean yellow mosaic India virus*	2671	India
LN846611	*Phaseolus vulgaris*	*Tomato yellow leaf curl Malaga virus*	2781	Morocco
KR052025	*Glycine max*	*Mungbean yellow mosaic India virus*	2746	India
KR071789	*Glycine max*	*Papaya leaf crumple virus*	2736	India
KP784661	*Lablab purpureus*	*Dolichos yellow mosaic virus*	2760	India
KP784665	*Vigna mungo*	*Mungbean yellow mosaic virus*	1164	India
KJ624994	*Phaseolus vulgaris*	*Squash leaf curl virus*	2667	Egypt
KP455992	*Vigna mungo*	*Mungbean yellow mosaic virus*	2731	India
KP677496	*Vigna mungo*	*Mungbean yellow mosaic India virus*	2504	India
KP313758	*Vigna radiata*	*Mungbean yellow mosaic India virus*	2746	India
LC010949	*Samanea saman*	*Bhendi yellow vein mosaic virus*	667	India
KP319016	*Vigna radiata*	*Mungbean yellow mosaic virus*	2649	India

KP319017	*Vigna radiata*	*Mungbean yellow mosaic virus*	2656	India
KJ585666	*Phaseolus vulgaris*	*Tomato yellow leaf curl virus*	2774	Japan
KJ939707	*Phaseolus lunatus*	*Bean golden mosaic virus*	2616	Brazil
KJ939766	*Macroptilium lathyroides*	*Bean golden mosaic virus*	2615	Brazil
KJ939779	*Phaseolus vulgaris*	*Bean golden mosaic virus*	2617	Brazil
KJ939854	*Phaseolus vulgaris*	*Macroptilium yellow spot virus*	2649	Brazil
KJ939856	*Macroptilium lathyroides*	*Macroptilium yellow spot virus*	2650	Brazil
KJ939858	*Phaseolus vulgaris*	*Macroptilium yellow spot virus*	2648	Brazil
KJ939895	*Macroptilium lathyroides*	*Macroptilium yellow spot virus*	2658	Brazil
KJ939898	*Macroptilium lathyroides*	*Macroptilium yellow vein virus*	2623	Brazil
KJ939916	*Macroptilium lathyroides*	*Soybean chlorotic spot virus*	2623	Brazil
KJ747961	*Vigna mungo*	*Mungbean yellow mosaic India virus*	920	India
KJ747962	*Vigna mungo*	*Mungbean yellow mosaic India virus*	920	India
KJ756846	*Glycine max*	*Mungbean yellow mosaic India virus*	761	India
KJ481204	*Lablab purpureus*	*Dolichos yellow mosaic virus*	2761	India
KC790019	*Phaseolus vulgaris*	*Pepper leafroll virus*	992	Peru
KJ131001	*Pongamia pinnata*	*Mungbean yellow mosaic virus*	775	India
KF976401	*Vigna mungo*	*Mungbean yellow mosaic India virus*	991	India
KF947526	*Vigna mungo*	*Mungbean yellow mosaic virus*	2635	India
KF928962	*Vigna mungo*	*Mungbean yellow mosaic virus*	2668	India
KF444467	*Phaseolus vulgaris*	*Tomato leaf curl Sudan virus*	2779	Saudi Arabia
KF440686	*Phaseolus vulgaris*	*Tomato leaf curl Gujarat virus*	2757	India
JX178666	*Crotalaria* sp.	*Tomato yellow spot virus*	239	Brazil
JQ733557	*Phaseolus vulgaris*	*Tobacco curly shoot virus*	2746	India
KC911717	*Vigna mungo*	*Mungbean yellow mosaic virus*	2728	India
KC911718	*Vigna mungo*	*Mungbean yellow mosaic virus*	2724	India
KC911719	*Vigna mungo*	*Mungbean yellow mosaic India virus*	2746	India
KC911721	*Vigna mungo*	*Mungbean yellow mosaic virus*	2734	India
KC911722	*Vigna mungo*	*Mungbean yellow mosaic virus*	2728	India
KC911723	*Vigna mungo*	*Mungbean yellow mosaic virus*	2738	India
KC911724	*Vigna mungo*	*Mungbean yellow mosaic virus*	2653	India

KC911731	*Vigna mungo*	*Mungbean yellow mosaic virus*	2651	India
KC852201	*Glycine max*	*Mungbean yellow mosaic India virus*	741	India
KC852202	*Lablab purpureus*	*Dolichos yellow mosaic virus*	471	India
KC852203	*Glycine max*	*Mungbean yellow mosaic India virus*	277	India
KC852204	*Glycine max*	*Mungbean yellow mosaic India virus*	2750	India
KF218188	*Cyamopsis tetragonoloba*	*Radish leaf curl virus*	2732	India
KC836518	*Glycine max*	*Mungbean yellow mosaic India virus*	1089	India
KC684949	*Vigna unguiculata*	*Tomato yellow leaf curl virus*	281	China
KC683813	*Vigna mungo*	*Mungbean yellow mosaic virus*	289	India
KC683814	*Glycine max*	*Mungbean yellow mosaic virus*	655	India
KC683816	*Vigna mungo*	*Mungbean yellow mosaic virus*	655	India
JX415184	*Crotalaria* sp.	*Euphorbia yellow mosaic virus*	2611	Brazil
JX415190	*Crotalaria juncea*	*Tomato severe rugose virus*	2591	Brazil
JX415191	*Crotalaria juncea*	*Euphorbia yellow mosaic virus*	2611	Brazil
KC854405	*Cyamopsis tetragonoloba*	*Cyamopsis tetragonoloba leaf curl virus*	1353	India
KC508642	*Phaseolus lunatus*	*Soybean chlorotic blotch virus*	2722	Nigeria
HF922628	*Glycine max*	*Mungbean yellow mosaic India virus*	2745	India
KC432613	*Phaseolus vulgaris*	*Tomato yellow leaf curl virus*	548	Saudi Arabia
KC432615	*Phaseolus vulgaris*	*Tomato yellow leaf curl virus*	548	Saudi Arabia
KC243784	*Vigna mungo*	*Mungbean yellow mosaic India virus*	920	India
KC243785	*Vigna mungo*	*Mungbean yellow mosaic India virus*	920	India
KC019303	*Phaseolus vulgaris*	*Mungbean yellow mosaic India virus*	2738	India
KC019307	*Phaseolus vulgaris*	*Horsegram yellow mosaic virus*	2677	India
JQ283246	*Phaseolus vulgaris*	*Bean chlorotic mosaic virus*	2566	Venezuela
JX997780	*Vigna unguiculata*	*Tomato yellow leaf curl virus*	526	China
KC004091	*Phaseolus vulgaris*	*Macroptilium yellow spot virus*	2649	Brazil
KC004103	*Macroptilium lathyroides*	*Macroptilium yellow spot virus*	2658	Brazil
KC004111	*Phaseolus vulgaris*	*Macroptilium yellow spot virus*	2658	Brazil
KC004112	*Macroptilium lathyroides*	*Macroptilium yellow spot virus*	2650	Brazil
KC004113	*Phaseolus vulgaris*	*Macroptilium yellow spot virus*	2650	Brazil
KC004132	*Macroptilium lathyroides*	*Macroptilium yellow spot virus*	2649	Brazil

KC004133	*Macroptilium lathyroides*	*Macroptilium yellow spot virus*	2648	Brazil
KC004134	*Phaseolus vulgaris*	*Macroptilium yellow spot virus*	2650	Brazil
JX480481	*Medicago polymorpha*	*Watermelon chlorotic stunt virus*	655	Iran
JX480483	*Melilotus indicus*	*Watermelon chlorotic stunt virus*	655	Iran
JX480485	*Melilotus dentatus*	*Watermelon chlorotic stunt virus*	655	Iran
JQ269833	*Vigna mungo*	*Mungbean yellow mosaic India virus*	540	India
JX436473	*Trigonella foenum-*	*Ageratum enation virus*	2753	India
JX315325	*Dolichos*	*Dolichos yellow mosaic virus*	2761	India
JX258325	*Cajanus cajan*	*Rhynchosia mosaic virus*	748	India
JX244172	*Vigna radiata*	*Mungbean yellow mosaic virus*	2730	Viet Nam
JX122965	*Glycine max*	*Soybean chlorotic spot virus*	2623	Brazil
JX110618	*Vigna mungo*	*Mungbean yellow mosaic India virus*	2746	India
HE861940	*Glycine max*	*Ageratum enation virus*	2752	India
JQ654460	*Vigna radiata var. radiata*	*Tomato leaf curl Joydebpur virus*	2769	India
JN543395	*Phaseolus vulgaris*	*Mungbean yellow mosaic India virus*	2729	Nepal
JN543396	*Phaseolus vulgaris*	*Mungbean yellow mosaic India virus*	2668	Nepal
JQ398670	*Vigna mungo*	*Mungbean yellow mosaic virus*	2679	India
JQ326957	*Phaseolus vulgaris*	*Tomato yellow leaf curl virus*	2781	China
JN419000	*M. atropurpureum*	*Euphorbia yellow mosaic virus*	2609	Brazil
JN419003	*Macroptilium lathyroides*	*Bean golden mosaic virus*	2614	Brazil
JN419005	*Macroptilium lathyroides*	*Macroptilium yellow spot virus*	2658	Brazil
JN419010	*M. atropurpureum*	*Tomato crinkle leaf yellows virus*	2586	Brazil
JN419015	*Calopogonium*	*Macroptilium yellow spot virus*	2649	Brazil
JN419018	*Macroptilium lathyroides*	*Macroptilium yellow spot virus*	2649	Brazil
JN419019	*Canavalia sp.*	*Macroptilium yellow spot virus*	2658	Brazil
JN419020	*Macroptilium lathyroides*	*Macroptilium yellow spot virus*	2658	Brazil
JN419021	*Macroptilium* sp.	*Macroptilium yellow vein virus*	2656	Brazil
JQ004982	*Vigna radiata*	*Mungbean yellow mosaic virus*	1285	India
JF501726	*Phaseolus vulgaris*	*Tomato leaf curl Palampur virus*	2756	Iran
JN896941	*Vigna mungo*	*Ageratum enation virus*	771	India
JN847431	*Vigna radiata*	*Mungbean yellow mosaic virus*	238	India

JN807765	*Glycine max*	*Papaya leaf curl virus*	2746	India
JN850955	*Cyamopsis tetragonoloba*	*Squash leaf curl China virus*	453	India
JN609292	*Erythrina*	*Pepper golden mosaic virus*	1160	Costa Rica
JN887122	*Cyamopsis tetragonoloba*	*Tomato leaf curl New Delhi virus*	453	India
JN998448	*Cyamopsis tetragonoloba*	*Cyamopsis tetragonoloba leaf curl Sikar*	524	India
JN368432	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	2746	Indonesia
JN368438	*Glycine max*	*Mungbean yellow mosaic India virus*	2746	Indonesia
JN368440	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	2652	Indonesia
JN368446	*Glycine max*	*Mungbean yellow mosaic India virus*	2652	Indonesia
JN368447	*Glycine max*	*Mungbean yellow mosaic India virus*	2653	Indonesia
JF694449	*Phaseolus* sp.	*Bean golden mosaic virus*	2616	Brazil
JF694460	*Phaseolus* sp.	*Cleome leaf crumple virus*	2582	Brazil
JN181003	*Vigna mungo*	*Mungbean yellow mosaic India virus*	920	India
JN181006	*Vigna mungo*	*Mungbean yellow mosaic India virus*	920	India
FR682601	*Lotus edulis*	*Tomato yellow leaf curl virus*	347	Cyprus
FR682602	*Lotus halophilus*	*Tomato yellow leaf curl virus*	347	Cyprus
FR682603	*Scorpiurus muricatus*	*Tomato yellow leaf curl virus*	347	Cyprus
HQ912025	*Vigna radiata var. radiata*	*Mungbean yellow mosaic India virus*	565	India
GU434143	*Vigna unguiculata subsp.*	*Tomato yellow leaf curl virus*	2781	China
FR837935	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	2746	Pakistan
HQ876467	*Mimosa pudica*	*Mimosa yellow vein virus*	774	India
HQ162271	*Glycine max*	*Kudzu mosaic virus*	2731	Viet Nam
HQ221570	*Vigna mungo*	*Mungbean yellow mosaic India virus*	575	India
FR714861	*Phaseolus vulgaris*	*Mungbean yellow mosaic India virus*	2672	India
FN794200	*Phaseolus vulgaris*	*Mungbean yellow mosaic India virus*	2746	India
HM357459	*Phaseolus vulgaris*	*Sida micrantha mosaic virus*	2692	Brazil
GQ200446	*Crotalaria juncea*	*Papaya leaf curl virus*	2738	India
GU320573	*Phaseolus vulgaris*	*Bean mosaic crinkle Barbados virus*	1314	Barbados
GU320574	*Macroptilium* sp.	*Bean mosaic crinkle Barbados virus*	1462	Barbados
GQ472985	*Glycine max*	*Soybean chlorotic blotch virus*	2708	Nigeria
GQ472988	*Centrosema pubescens*	*Soybean chlorotic blotch virus*	2647	Nigeria

AM992617	*Vigna radiata*	*Mungbean yellow mosaic India virus*	2672	Pakistan
AM950268	*Vigna radiata*	*Mungbean yellow mosaic India virus*	2746	Pakistan
GU585391	*Trigonella* sp.	*Tomato yellow leaf curl virus*	608	Iran
GU585392	*Melilotus officinalis*	*Tomato yellow leaf curl virus*	608	Iran
GU591170	*Vigna unguiculata*	*Dolichos yellow mosaic virus*	774	India
GU591171	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	774	India
AM948961	*Glycine max*	*Pedilanthus leaf curl virus*	2760	Pakistan
GU385879	*Cyamopsis tetragonoloba*	*Cotton leaf curl Kokhran virus*	2736	India
GU385766	*Crotalaria juncea*	*Cotton leaf curl virus*	854	India
FM958506	*Vigna radiata*	*Mungbean yellow mosaic India virus*	2660	Pakistan
FN645915	*Cyamopsis tetragonoloba*	*Croton yellow vein mosaic virus*	2755	India
FN645916	*Cyamopsis tetragonoloba*	*Tomato leaf curl New Delhi virus*	1504	India
GU323321	*Phaseolus vulgaris*	*Horsegram yellow mosaic virus*	2735	Sri Lanka
FM955601	*Rhynchosia capitata*	*Papaya leaf curl virus*	2754	Pakistan
FN435274	*Neustanthus phaseoloides*	*African cassava mosaic virus*	2779	Democratic
FN435276	*Centrosema pubescens*	*African cassava mosaic virus*	2779	Democratic
FN435278	*Neustanthus phaseoloides*	*East African cassava mosaic virus*	2799	Democratic
FN435280	*Centrosema pubescens*	*East African cassava mosaic virus*	2799	Democratic
FM242702	*Rhynchosia capitata*	*Mungbean yellow mosaic virus*	2676	Pakistan
FN543425	*Mucuna* sp.	*Velvet bean severe mosaic virus*	2733	India
AM999981	*Rhynchosia minima*	*Rhynchosia yellow mosaic virus*	2740	Pakistan
GQ387501	*Vigna radiata*	*Mungbean yellow mosaic India virus*	774	India
FM208833	*Glycine max*	*Mungbean yellow mosaic India virus*	2746	Pakistan
FM208834	*Glycine max*	*Mungbean yellow mosaic India virus*	2746	Pakistan
FM208835	*Vigna mungo*	*Mungbean yellow mosaic India virus*	2751	Pakistan
FM208840	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	2746	Pakistan
FM208841	*Vigna mungo*	*Mungbean yellow mosaic India virus*	2746	Pakistan
FM208842	*Vigna radiata*	*Mungbean yellow mosaic India virus*	2749	Pakistan
FM208845	*Vigna mungo*	*Mungbean yellow mosaic India virus*	2745	Pakistan
FM208846	*Vigna radiata*	*Mungbean yellow mosaic India virus*	2746	Pakistan
GQ357649	*Phaseolus vulgaris*	*Sida golden mosaic virus*	2633	USA

GQ352453	*Desmodium* sp.	*Rhynchosia golden mosaic Yucatan virus*	2601	Mexico
FM161881	*Glycine max*	*Mungbean yellow mosaic India virus*	2674	Pakistan
FJ686693	*Glycine max*	*Sida micrantha mosaic virus*	2669	Brazil
FJ821189	*Vigna silvestris*	*Mungbean yellow mosaic India virus*	1541	
FJ792608	*Rhynchosia minima*	*Rhynchosia golden mosaic Yucatan virus*	2542	Mexico
EU523045	*Glycine max*	*Mungbean yellow mosaic India virus*	2747	India
FJ687389	*Glycine max*	*Tomato yellow leaf curl Thailand virus*	771	Thailand
FJ686696	*Glycine max*	*Okra mottle virus*	2652	Brazil
FJ663015	*Vigna silvestris*	*Mungbean yellow mosaic virus*	774	India
FJ665283	*Glycine max*	*Bean golden mosaic virus*	2611	Brazil
EU367501	*Glycine max*	*Soybean mottle mosaic virus*	438	Nigeria
FJ595240	*Vigna radiata*	*Mungbean yellow mosaic virus*	1089	India
FJ538207	*Fabaceae*	*Tomato yellow spot virus*	2632	Argentina
FJ539014	*Pueraria montana var.*	*Kudzu mosaic virus*	2729	China
FJ513076	*Phaseolus vulgaris*	*Sida golden mosaic virus*	1112	USA
AM932425	*Phaseolus vulgaris*	*Horsegram yellow mosaic virus*	2735	India
AM932427	*Vigna unguiculata*	*Horsegram yellow mosaic virus*	2735	India
AM932429	*Phaseolus lunatus*	*Horsegram yellow mosaic virus*	2734	India
AM932430	*Phaseolus lunatus*	*Horsegram yellow mosaic virus*	2670	India
EU685320	*Leucaena leucocephala*	*African cassava mosaic virus*	2781	Nigeria
EU685321	*Leucaena leucocephala*	*East African cassava mosaic Cameroon*	2800	Nigeria
EU685322	*Senna occidentalis*	*African cassava mosaic virus*	2781	Nigeria
EU685323	*Senna occidentalis*	*East African cassava mosaic Cameroon*	2800	Nigeria
EU685325	*Glycine max*	*African cassava mosaic virus*	2781	Nigeria
EU709521	*Phaseolus vulgaris*	*Merremia mosaic virus*	1291	Belize
EU367500	*Glycine max*	*African cassava mosaic virus*	987	Nigeria
EU339936	*Glycine max*	*Rhynchosia golden mosaic virus*	2604	Mexico
EU339938	*Rhynchosia minima*	*Rhynchosia golden mosaic virus*	2605	Mexico
DQ836302	*Fabaceae*	*Bean golden yellow mosaic virus*	756	Mexico
AM701757	*Phaseolus vulgaris*	*Bean leaf curl Madagascar virus*	2754	Madagascar
EF601701	*Phaseolus vulgaris*	*Tomato yellow leaf curl virus*	536	Guadeloupe

EF541485	*Macroptilium lathyroides*	*Bean golden mosaic virus*	1011	Brazil
DQ641690	*Pueraria montana*	*Kudzu mosaic virus*	2731	Viet Nam
EF492517	*Calopogonium mucunoides*	*Calopogonium mucunoides begomovirus - Trinidad*	204	Trinidad and Tobago
EF119337	*Crotalaria juncea*	*Crotalaria juncea begomovirus*	897	India
AM418398	*Phaseolus vulgaris*	*Tomato yellow leaf curl virus*	578	Greece
DQ318934	*Rhynchosia minima*	*Tobacco apical stunt virus*	756	Mexico
AM157413	*Dolichos* sp.	*Dolichos yellow mosaic virus*	2762	India
DQ914441	*Fabaceae*	*China bean begomovirus*	522	China
DQ865203	*Vigna aconitifolia*	*Mungbean yellow mosaic virus*	2679	India
DQ515843	*Dolichos* sp.	*Mungbean yellow mosaic India virus*	1027	India
DQ515844	*Glycine max*	*Mungbean yellow mosaic India virus*	1028	India
DQ515845	*Dolichos* sp.	*Mungbean yellow mosaic India virus*	1024	India
DQ515846	*Vigna radiata*	*Mungbean yellow mosaic India virus*	1015	India
DQ515848	*Vigna mungo*	*Mungbean yellow mosaic India virus*	1016	India
DQ515849	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	1023	India
DQ515850	*Dolichos* sp.	*Mungbean yellow mosaic India virus*	1023	India
DQ515851	*Glycine max*	*Mungbean yellow mosaic India virus*	1027	India
DQ515852	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	1024	India
DQ515853	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	530	India
DQ515854	*Glycine max*	*Mungbean yellow mosaic India virus*	531	India
DQ515855	*Dolichos* sp.	*Mungbean yellow mosaic India virus*	531	India
DQ515856	*Dolichos* sp.	*Mungbean yellow mosaic India virus*	531	India
DQ515857	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	531	India
DQ515858	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	531	India
DQ515859	*Vigna mungo*	*Mungbean yellow mosaic India virus*	529	India
DQ515860	*Vigna radiata*	*Mungbean yellow mosaic India virus*	396	India
DQ515861	*Cyamopsis tetragonoloba*	*Mungbean yellow mosaic India virus*	1016	India
DQ389145	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	1122	India
DQ389146	*Glycine max*	*Mungbean yellow mosaic virus*	1122	India
DQ389147	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	1121	India
DQ389148	*Dolichos* sp.	*Mungbean yellow mosaic India virus*	1122	India

DQ389149	*Cassia* sp.	*Mungbean yellow mosaic India virus*	1114	India
DQ389150	*Glycine max*	*Mungbean yellow mosaic India virus*	1108	India
DQ389151	*Glycine max*	*Mungbean yellow mosaic India virus*	1124	India
DQ389152	*Glycine max*	*Mungbean yellow mosaic India virus*	1122	India
DQ389154	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	2747	India
DQ389155	*Glycine max*	*Mungbean yellow mosaic India virus*	1122	India
DQ406672	*Rhynchosia minima*	*Rhynchosia golden mosaic Sinaloa virus*	2578	Mexico
AM233490	*Vigna aconitifolia*	*Mungbean yellow mosaic India virus*	771	Pakistan
AY939925	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	2669	India
DQ369046	*Fabaceae*	*Begomovirus isolate bean Iguala*	527	Mexico
DQ347945	*Glycine max*	*Chino del tomate virus*	2609	Mexico
DQ343283	*Glycine max*	*Cotton leaf curl Kokhran virus*	771	
DQ343284	*Glycine max*	*Tomato leaf curl virus*	771	
DQ343285	*Glycine max*	*Pepper leaf curl virus*	771	
DQ256460	*Phaseolus vulgaris*	*Tomato yellow leaf curl China virus*	2731	China
AM156857	*Glycine max*	*Soybean crinkle leaf virus*	1270	Thailand
AM156859	*Glycine max*	*Malvastrum yellow vein Yunnan virus*	1278	Thailand
AM156861	*Senna obtusifolia*	*Foetid cassia leaf curl virus-[Thailand]*	1310	Thailand
AY957561	*Glycine max*	*Rhynchosia golden mosaic virus*	915	Mexico
AY955102	*Leucaena multicapitula*	*Rhynchosia golden mosaic virus*	916	Mexico
AM087116	*Vigna unguiculata subsp.*	*Mungbean yellow mosaic India virus*	539	Bangladesh
AM087117	*Vigna unguiculata subsp.*	*Mungbean yellow mosaic India virus*	539	Bangladesh
DQ061273	*Lablab purpureus*	*Mungbean yellow mosaic India virus*	2665	India
AJ810372	*Vigna unguiculata*	*Tomato leaf curl virus*	771	India
AY905553	*Glycine max*	*Pepper golden mosaic virus*	571	Mexico
AJ867554	*Glycine max*	*Mungbean yellow mosaic virus*	2670	India
AY618902	*Vigna unguiculata*	*Cowpea golden mosaic virus*	2743	India
AY639879	*Senna tora*	*Tomato yellow leaf curl Thailand virus*	771	Thailand
AY547317	*Dolichos*	*Mungbean yellow mosaic India virus*	2736	India
AJ627904	*Macrotyloma uniflorum*	*Horsegram yellow mosaic virus*	2728	India
AY271892	*Vigna radiata var. radiata*	*Mungbean yellow mosaic virus*	2729	Cambodia

AY436328	*Glycine max*	*Sida mottle virus*	677	
AY269990	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	2751	Pakistan
AY269991	*Glycine max*	*Mungbean yellow mosaic virus*	2729	Pakistan
AY269992	*Vigna radiata*	*Mungbean yellow mosaic India virus*	2746	Pakistan
AY271895	*Vigna radiata*	*Mungbean yellow mosaic India virus*	2746	Nepal
AY271896	*Vigna radiata*	*Mungbean yellow mosaic virus*	2737	India
AY217344	*Rhynchosia minima*	*Rhynchosia minima Trinidad virus*	478	Trinidad and
AJ512495	*Vigna radiata*	*Mungbean yellow mosaic India virus*	2746	Pakistan
AF449192	*Fabaceae*	*Macroptilium mosaic Puerto Rico virus*	2615	USA
AY099065	*Vigna unguiculata*	*Cowpea golden mosaic virus*	160	
AF481865	*Vigna unguiculata*	*Mungbean yellow mosaic India virus*	2746	India
AF442117	*Rhynchosia minima*	*Rhynchosia mosaic virus*	750	USA
AY062025	*Cajanus cajan*	*Rhynchosia mosaic virus*	533	USA
AJ421642	*Glycine max*	*Mungbean yellow mosaic virus*	2729	India
AF439402	*Phaseolus vulgaris*	*Calopogonium golden mosaic virus*	1365	Costa Rica
AY049772	*Glycine max*	*Mungbean yellow mosaic India virus*	2748	India
AF239671	*Rhynchosia minima*	*Rhynchosia golden mosaic virus*	2624	Honduras
AF176092	*Macroptilium lathyroides*	*Macroptilium mosaic virus*	750	USA
AF188708	*Vigna unguiculata*	*Cowpea golden mosaic virus*	1365	
AF098940	*Macroptilium lathyroides*	*Macroptilium golden mosaic virus*	1405	Jamaica
U73511	*Macroptilium lathyroides*	*Macroptilium golden mosaic virus*	159	
U70386	*Macroptilium lathyroides*	*Sida golden mosaic virus*	151	
U00121	*Phaseolus vulgaris*	*Bean calico mosaic virus*	659	